## *The statistical lensing of QSOs*

Myers, Adam David

# The Statistical Lensing of QSOs

*Adam David Myers*

A thesis submitted to the University of Durham
in accordance with the regulations for admission to the
Degree of Doctor of Philosophy.



Department of Physics



University of Durham
2003

**Abstract: The Statistical Lensing of QSOs**

by **Adam David Myers**

We use the 2dF QSO Redshift Survey, to investigate whether QSOs are detectably gravitationally lensed. Lensing could magnify and distort light from QSOs, influencing QSO numbers near galaxies, which trace structure in our Universe.

Following Boyle, Fong & Shanks (1988), we find a $3\sigma$ anti-correlation between QSOs and galaxy groups of strength $\omega_{qg}(< 10') = -0.049$. We limit absorption by dust in groups to $A_B < 0.04$ mag. To explain the anti-correlation by dust would need $A_B \approx 0.2$ mag. We demonstrate that if the dearth of QSOs around groups is due to statistical lensing, more mass would be required in groups than $\Omega_m = 0.3$ models suggest.

We use a mock catalogue to test how many of our "2D" galaxy groups, which are detected using angular information, are associated in redshift-space. We then utilise 2dF Galaxy Redshift Survey groups, which are selected to trace dark matter haloes, to test the hypothesis that there is more mass in groups than $\Omega_m = 0.3$ models suggest, finding we cannot discount a lensing mass of 2dFGRS groups that is consistent with $\Lambda$CDM.

We find QSOs and galaxies are also anti-correlated at the $3\sigma$ level, with strength $\omega(< 10') = -0.007$ and use stars as a control sample to rule out observational systematics as a cause. By measuring QSO colours as a function of QSO-galaxy separation, we argue that obscuration by dust in galaxies could explain at most 30-40 per cent of the anti-correlation. We show that if the anti-correlation is due to lensing, galaxies would be anti-biased ($b \sim 0.05$) on small scales.

We discuss two surveys carried out to count faint QSOs, which newly identify 160 QSOs. We calculate that the faint-end QSO number-counts have a slope of $0.29 \pm 0.03$. Finally, we use our faint QSO data, to estimate that $\sim$ 85(75) per cent of $g < 21.15$ ($\geq 21.15$) candidates targeted by the 2dFSDSS survey will be QSOs.

# Contents

# List of Figures

# List of Tables

# Preface

*The Statistical Lensing of QSOs*

The work described in this thesis was undertaken between October 2000 and July 2003 whilst the author was a research student under the supervision of Prof. Tom Shanks in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at this (or any other) university.

Other than where directly referenced in the body of this thesis, all other content is the author's own work. The main collaborators involved in this thesis are Dr. P. J. Outram and Prof. T. Shanks.

Some of the results and analysis from this thesis have appeared in the following papers:

- Myers A. D., Outram P. J., Shanks T., Boyle B. J., Croom S. M., Loaring N. S., Miller L. & Smith R. J, 2003, MNRAS, 342, 467
- Outram P. J., Hoyle F., Shanks T., Croom S. M., Boyle B. J., Miller L., Smith, R. J. & Myers A. D., 2003, MNRAS, 342, 483

and the following conference proceedings.

- Myers A. D., Shanks T., Outram P. J. & The 2dF QSO Redshift Team, in A New Era in Cosmology (ASP Conference Proceedings), eds. T. Shanks & N. Metcalfe (2001)
- Outram P. J, Shanks T., Myers A. D., Boyle B. J., Croom S. M., Hoyle F., Loaring N., Miller L. & Smith R. J., in Galaxy Evolution: Theory and Observations, eds. V. Avila-Reese, C. Firmani, C. Frenk & C. Allen, RevMexAA SC (2002)

# Acknowledgements

Firstly, I must thank my Ph.D. supervisor, Tom Shanks, and his day-to-day supervisor Phil Outram for all of their ideas, suggestions and support over the last three years. Together, Tom and Phil ensured my time was interesting, challenging and filled with debate. I will also never forget the supervison of Dr. John Lucey in my undergraduate days - his enthusiasm for astronomy is probably what led me to study for this Ph.D. in the first place. Much of the computer-related workload of research for this thesis has been eased by Alan Lotts, Peter Draper and Nigel Metcalfe, and I have greatly appreciated their valuable time. Welcome scientific criticism and input has been provided by Scott Croom, Lance Miller, Shaun Cole and, no doubt, many other people whom I've forgotten. I am additionally indebted to Robert Smith and Brian Boyle for allowing me use of the 2dF QSO Redshift Survey before its publication. The "additionally" caveat in that last sentence was to save me having to thank Scott Croom, Tom Shanks and Lance Miller again. Thanks again. Thanks also to Vince Eke and the 2dF Galaxy Redshift Survey Collaboration for allowing me to study the 2dF galaxy group catalogue prior to its publication. I also appreciated Fiona Hoyle giving me an early copy of her thesis.

Postgraduate work at Durham has been made a great deal more pleasant by the people I've shared offices with over the last few years (they're also the people you tend to learn most of the day-to-day tricks from), so I'd like to thank Chris, Steve, Marc and Professor Sir Arnold Wolfendale. And Bill - a constant distraction - I will never agree with you that gauge theories are better reformulated by canonical loop representations than as spin foams but you might have a point about Wayne Rooney. I never shared an office with Geoff Busswell, which probably allowed both of us to do some work, as I did share just about every pool table in Durham with him.

I would like to thank my parents for their emotional and financial support over a difficult few years. My interest in studying the Universe was no doubt fuelled by my Dad's stories and taste in literature and my general interest in study by my Mum's enthusiasm. They're both at fault for my incessant work ethic. Although, my work ethic (and probably general intelligence) pales in comparison to my Granddad's. I'm also willing to bet that if any of my family make an ill-advised attempt to read this thesis, it'll be my Uncle Vic - turn back now, wary traveller.

I have been lucky to have a number of excellent teachers and lecturers over my last twenty-one years of study. Some of them inspired, some of them are just very talented teachers. In no particular order, I'd like to mention Prof. B.K. Tanner, Dr. S. Brand, Prof. E. W. N. Glover, Dr. R. G. Bower, Dr. C. S. Adams, Dr. C. J. Maxwell, Dr. A. Signer, Tom Wilson, Ron Wiggins and Ron Mathieson, whose lectures, lessons and/or tutorials I have immensely enjoyed.

As I'm beginning to feel like the stereotypical Oscar winner with 30 seconds to thank the Academy.....and so forth, here's a list of great people: Phil Outram, who has great taste in hiking, whisky and politics but bad taste in football teams; Gav "Gaz" Zephir, who likes Christmas music; Jim Taylor, who could go on an apple-break for England; Chris "Dutch" Nillesen, inter-planetary good vibe zone; Andy Richmond, who is God's great solipsistic experiment; Jeff "Dobsy" Shuttleworth, who's asleep; Chris "Ian" Rowe, who broke the levitating clock; and Jennie B. Morris, my 1000th paper crane. A less verbose list of great people who I've enjoyed playing or watching football with: Mike Daniel, Marc Vallbe, Nigel Metcalfe, Phil Outram, J. William Frith, Graham Smith, Carlton Baugh, Shaun Cole, Ben Moore, Marc Swinbank, Josè Cruz-da-angela, Clive Doloughan, George Teasdale, Ian Manfren and especially Wayne Dobby and Chris Power and Peder Norberg and Richard McDermid.

Finally - my unbelievable Jenny. The lens through which I view my universe.

**pity this busy monster,manunkind,**

not. Progress is a comfortable disease:
your victum(death and life safely beyond)

plays with the bigness of his littleness
-electrons deify one razorblade
into a mountainrange;lenses extend

unwish through curving wherewhen until unwish
returns on its unself.

A world of made
is not a world of born-pity poor flesh

and trees,poor stars and stones,but never this
fine specimen of hypermagical

ultraomnipotence. We doctors know

a hopeless case if-listen:there's a hell
of a good universe next door;let's go

**e. e. cummings, 1944**

# Chapter 1 | *Introduction*

Some of the earliest written records of human civilisation detail the stars above. Trace the dome of the night sky, empyreal, inspirational, filled with burning matter, arcing overhead. Texts in Ancient Chinese, Ancient Babylonian, in Ancient Hebrew talk of the celestial vault, of the columns that support the sky. Over two thousand years ago the Greeks wondered at the stuff of the Universe, debated at the form of the firmament. What supports and shapes the heavens? What moulds the light that arrows through the scattered wash of midnight? Questions of shape, of structure, of the intertwining of darkness and light. Philosophies of the Universal matter.

And though the landscape expanded, though the technology with which to gaze at the sky and sample the Cosmos progressed, the fundamental questions did not change. In the millenia since the Ancients, keen minds considered the geometry of our galaxy and the composition of the stars. The shape and the matter. At the dawn of the twentieth century, in an era driven by Albert Einstein's belief in the equivalence of gravitational and inertial acceleration, the shape and the matter were married. In a relativistic world, geometry tells matter how to move and matter tells geometry how to curve (Misner, Thorne & Wheeler 1973). And the fundamental questions were still questions of the shape and of the stuff of the Universe.

Understanding of our Universe expanded rapidly through observations made in the early twentieth century, as did the theoretical size of the Universe, fuelled by astounding advances in technology. Spectrographs that could capture the rainbow of light from bright stars became prevalent. Telescopes were built with 1.5-metre mirrors, even, staggeringly, with 2.5-metre mirrors. Through the first three decades of the Twentieth Century, work on "spiral nebulae" and Cepheid variable stars by Slipher and Shapley culminated with Edwin Hubble making the remarkable observations that galaxies are immensely distant and, in the vast majority of cases, moving quickly away from us. Complementary theory emerge, allowing Einstein, de Sitter, Friedmann, Lemaitre, Robertson and Walker to ultimately provide us with a world model of an expanding Universe where matter and geometry entangle. The discipline of cosmology was born and the foundations to study



the Universe on yet vaster scales laid down. Like the ancients, like Slipher and Shapley and Hubble, observational cosmologists of the mid-twentieth century would look at the distribution of luminous sources in the night and they would ask questions of geometry, of how light traces the matter distribution, of the stuff of our Universe.

And though the landscape has expanded, though the technology with which to gaze at the sky and sample the Cosmos has progressed, still the fundamental questions have not essentially changed.

## 1.1 The Major Parameters of The Cosmological World Model

In the 1920s, as Einstein's General Theory of Relativity became the prominent theory of gravity, matter and geometry, the principal tenet underlining cosmological thought was the Perfect Cosmological Principle, the idea that integrated over large enough timescales, over large enough angles and over large enough distances, the Universe must appear homogenous. To Einstein's dismay, however, his inital formulation of Relativity allowed no static solutions for the fabric of space-time. His theory lay in contradiction to the Perfect Cosmological Principle, the universe of his General Relativity was a malleable balloon, ever-deflating, or ever-expanding (Einstein 1916). In relief, Einstein noted that the addition of a "cosmological" constant to his equations allowed for a static universe (Einstein 1917), as the astronomical community expected. A second world model, proposed by de Sitter (de Sitter 1916) had no cosmological constant but was quasi-static as time measurements in this universe depended on the observer's position. Curiously, light emitted by particles placed in de Sitter's cosmology would appear "spectrally shifted" to different observers, at turns, both bluewards and redwards. The Einstein and de Sitter models represented a miniscule fraction of a multitude of dynamic cosmologies (Friedmann 1923, Lemaitre 1927).

The work of Edwin Hubble modified the cornerstone of cosmology. In the space of five years, Hubble published papers showing, firstly, that the light from nearly all galaxies is "red-shifted" upon observation (Hubble 1929) and that the numbers of galaxies appear to sample the Universe in a manner that is homogenous (Hubble 1934). Robertson (1935) and Walker (1936) formulated a description of the Universe that embodied the Cosmological Principle, that the Universe is homogenous and isotropic while allowing Friedmann and Lemaitre's perceptions of a dynamic universe, changing in time. In a universe described by the metric of Robertson and Walker, particles are distributed with

homogeneity, the light they emit is redshifted and the universe is ever-expanding. There are a plethora of acceptable universes, embodied by Friedmann's equations; where the geometry of the Universe is inextricably linked to its content. In Appendix A we derive the major cosmological parameters, the standard descriptors of our Universe, from the Roberson-Walker Metric and the Friedmann Equations.

## 1.2 Modern Cosmology

The 1960s brought two events that would have a profound influence on cosmology, driven by the advent of radio astronomy in the 1940s and 1950s. The first redshift of a quasi-stellar radio source (QSO) was measured (Schmidt 1963), implying that QSOs were luminous enough to probe the Universe at redshifts ten times higher than probed by galaxies, at a time when the Universe was 20 per cent of its current age (Schmidt 1965). Such an impressive look-back time was dwarfed by the discovery (Penzias & Wilson 1965) of the Cosmic Microwave Background (CMB) radiation, the black-body remnant of a time near the dawn of the expansion of the Cosmos, when the Universe was smaller, denser and hotter. The isotropy and homogeneity of the CMB implied homogeneity and isotropy on large scales, ratifying the Cosmological Principle. The standard model of the day was born, and it was widely accepted that the Universe was born in a "Big Bang"; the simultaneous emergence of matter and radiation throughout the interstices of the Robertson-Walker metric in an agglomeration that cooled as the Universe expanded, freezing out matter in a manner that garners an excellent prediction of the relative abundances of the elements (Peebles 1966).

Though the Big Bang Model says much, much of modern cosmology has been driven by the questions that the model can't address. Particularly, in a Universe that is largely homogenous at its inception, how can the myriad structure seen in the clustering of galaxies and QSOs be explained? Through the 1970s, aided by the development of computers, standard methods of quantifying the clustering of extragalactic sources were developed and applied (e.g., Peebles 1973). The main statistics used to describe the distribution of extragalactic objects, the auto-correlation function and the power spectrum, remain the statistical tools of choice today. Essentially, both the (two-point) auto-correlation function and the power spectrum measure the mean density of sources at a given distance from any source, though the correlation function does so in real space and the power spectrum does so in Fourier space (see Appendix B).

Through the 1980s, the advent of inflationary theory (Guth 1981) provided an explanation of the large-scale uniformity of our Universe. A widely accepted picture of the evolution of structure begins with quantum fluctuations in a hot primeval universe, a universe that expands phenomenally rapidly. As the inflation slows, matter begins to collapse under the influence of gravity, seeding structure in the universe. The universe cools as it expands, and the energetic scattering reactions that were coupling photons and baryons eventually cease. Photons are freed to traverse the universe, allowing us to observe the cooling embers of the surface of last scattering as the CMB. On large scales, because of the finite speed of light, matter does not have time to gravitationally interact before the present day, explaining the Cosmological Principle. On smaller scales, we should see where photons coupled to large-scale structure traced the wells of gravity. A strong prediction of inflationary theory is that after the initial expansion of the universe, fluctuations in the density of the universe around the mean density should have a randomised Gaussian distribution. Such Gaussian fluctuations are completely characterised by their power spectrum, often called the intitial power spectrum. Thus, observations of the power spectrum of the CMB, or of QSO or galaxy clustering can powerfully probe cosmology, particularly, on moderate scales, the total mass of the Universe (embodied in $\Omega_m$, the fraction of the closure density of the Universe that is in matter) and the scale of the Universe (embodied in Hubble's Constant, $H = 100h\,\mathrm{km\,s^{-1}\,Mpc}$. The power spectrum is often characterised by the shape parameter $\Gamma = \Omega_m h$. Observing the power spectrum in structure or the CMB also constrains the form of the initial power spectrum.

It has long been known that the rotational velocities of galaxies as a function of radius do not drop off with the light, that, assuming Newtonian gravity is valid, there must be a great deal of mass in galaxies that is not luminous (Zwicky 1933). In the last twenty years, as knowledge of the initial spectrum and cosmological parameters has grown, simulations of the Universe have become sophisticated. These simulations must theorise the properties of the dark matter, which only seems to interact with baryonic matter through gravitation, and evolve them through gravitational interaction. The most succesful paradigm of the last few decades, collisionless Cold Dark Matter (e.g. Peebles 1982, Bond & Szalay 1983), suggests that dark matter is not energetic enough to venture far from inhomogeneities in the density field, collapsing to form massive rotating "haloes". As time progresses, smaller haloes merge to form larger haloes and baryons are trapped in these agglomerations of mass. Baryonic interaction is complicated as, unlike the dark matter, baryons collide and interact, they shock and heat each other, and eventually stars

form, allowing us to see the structure seeded by the Cold Dark Matter (CDM).

Exactly how the baryons trace the CDM is complicated. A widely adopted theory suggests baryons are "switched on" in areas of the density field which peak above a certain threshold value (Kaiser 1984). If the density field of the dark matter is Gaussian, this suggests that the bias can be parametrised by a single parameter, $b$, which measures the amplitude of clustering of dark matter compared to the amplitude of clustering of galaxies. The amplitude of clustering of dark matter, $\sigma_8$, has historically been measured in simulations on scales of $8\,h^{-1}$Mpc. In Appendix B, we provide brief mathematical descriptions of the power spectrum, the correlation function and the bias parameter.

In the last decade, many of the key cosmological parameters have been well constrained by large surveys. We have come full circle in the story of cosmology, in the sense that Einstein's original innovation, the cosmological constant, $\Lambda$, is now widely accepted as the mechanism by which observations are made consistent with each other and with the expectations of theory. There have been many recent observational highlights. The COBE satellite demonstrated the impressive uniformity of the CMB, while finding anisotropies in the temperature of the CMB of the order of $10^{-5}$K. The Supernova Cosmology Project (Perlmutter et al. 1999) have constructed a statistically significant set of observations of supernovae. Assuming supernovae have an intrinsic luminosity that is unaltered by their redshift, the apparent magnitudes and redshifts of Supernova Cosmology Project sources constrain the geometry of the Universe to the range $0.8\Omega_{\rm m} - 0.6\Omega_\Lambda \sim -0.2 \pm 0.1$. Measurements of the power spectrum of both galaxies and QSOs in the 2dF Redshift Surveys (Percival et al. 2001, Outram et al. 2003) have constrained the shape parameter to $\pm$ 25 per cent of $\Gamma = \Omega_m h = 0.2$. Key Project measurements of Hubble's Constant (Freedman et al. 2001) suggest $h = 0.72 \pm 0.08$. In the last year, the WMAP satellite has measured the power spectrum of the CMB to a resolution of about 20 arcminutes (around 20 times higher resolution than COBE), confirming the Gaussianity of fluctuations in the CMB (Komatsu et al. 2003) and independently constraining cosmological parameters with unprecedented accuracy ($h = 0.72 \pm 0.05, \Omega_m h^2 = 0.14 \pm 0.02, \sigma_8 = 0.9 \pm 0.1$). Probably any two of the above observations (consider Efstathiou 2002) and certainly any three are sufficient to say with some confidence that the geometry of our Universe is flat, with $\Omega_{\rm m} \sim 0.3$ and $\Omega_\Lambda \sim 0.7$. This widely accepted cosmology, is generally called the "Concordance Cosmology". Taken together with the Cold Dark Matter paradigm, it is often called $\Lambda$CDM .

## 1.3 Gravitational Lensing

Though the very existence of the field of cosmology itself tests the ideas of General Relativity, the bending of light by massive bodies is a more direct test. In the conclusion of his treatise *Opticks* in 1704, Newton qualitatively suggested that the path of a light ray might be lensed by a massive object and, further, that the effect should be greatest at the closest approach of the light to the mass. In 1804, Johann Soldner used Newtonian theory to calculate the deflection of a light-ray by its closest possible approach to the solar surface (Wambsganss 1998). Newtonian gravitation predicts that a close approach to the solar surface would deviate a light-beam by an angle of about $\alpha_\odot \sim 2GM_\odot/c^2R_\odot$. Initially, Einstein concurred with this figure (Einstein 1911) but as his formulation of General Relativity progressed, upgraded the prediction to $\alpha_\odot \sim 4GM_\odot/c^2R_\odot$ (Einstein 1916). That the prediction of General Relativity is correct is now established to hundredths of a per cent (Lebach et al. 1995).

The power of Einstein's prediction of the angle that light would be deflected by the sun is in its inversion. If we had no measure of the mass of the sun, we could use the angle with which the sun deflected background light to predict its mass. Weighing distant objects by the gravitational lensing of light is now an established tool of observational cosmology. Loosely, the field falls into two categories, strong lensing, which deals with light passing closely enough to a massive body that the ray is split along several lines of sight, and weak lensing, where the small-scale manipulation by foreground mass of the light from many background sources is analysed en masse. Since the discovery of the first multiply-imaged QSO (Walsh, Carswell & Weymann 1979), dozens of such systems have been discovered (Kochanek et al. 2002). Strong lensing is now a multi-faceted field, Einstein Rings, where an extended background source is split to form a ring around a foreground cluster, have been discovered (Hewitt et al. 1988), as have the giant arcs of background galaxies spattered around the periphery of the Einstein Ring in massive foreground clusters (Lynds & Petrosian 1986, Soucail et al. 1987). In the weaker regime, merely the alignments of background galaxies around a cluster can be used to trace the lensing mass (Tyson, Valdes & Wenk 1990), as background galaxies are "sheared" to align preferentially along radial lines plotted from the cluster centre.

Gravitational lensing is often pictured as the bending of light by the gravitational field of a massive body, however, as different elements of the cross-section of a light ray are in different parts of the field, a light ray can be tidally smeared or focussed by

lensing. Conservation of energy then implies a magnification (or de-magnification) of background sources by a foreground distribution of masses. Depending on the intrinsic form of the distribution of luminosites of a population of sources, weak lensing can thus change the number of sources counted near a foreground mass distribution. The first use of QSO numbers to probe this magnification bias was probably the study of Webster et al. (1988). In the decade or so since, the effect of magnification bias on the numbers of QSOs around foreground objects has been reconsidered several times but usually using only small samples of QSOs - and results have been contradictory or inconclusive. In Appendix C we discuss gravitational lensing theory and mathematically derive results that will be used in this thesis.

## 1.4 Thesis Aims

The 2dF QSO Redshift Survey (Croom et al. 2003) is the largest extant sample of spectroscopically identified QSOs. We propose to test the idea that the foreground mass distribution influences QSO numbers by the process of magnification bias, and use the extent of any effect to probe both modern and classic cosmological parameters.

In Chapter 2 we introduce the 2dF QSO Redshift Survey (2QZ). We discuss the production of masks used to mimic the fluctuating selection of objects in the 2QZ due to the variable completeness of the survey. We also derive results on the number-magnitude and number-redshift distributions of 2QZ QSOs, which will be repeatedly used in this thesis.

In Chapter 3, we reexamine the result of Boyle, Fong & Shanks (1988) that faint UVX-selected objects (i.e. QSOs) preferentially avoid foreground clusters, by cross-correlating 2QZ QSOs and groups of galaxies objectively selected from galaxy surveys in the region of the 2QZ. We discuss the effect that dust in galaxy groups could have on the relative number of QSOs around galaxy groups and pay close attention to any survey selection effects that could influence QSO numbers around groups. Finally, we discuss the expected influence of gravitational lensing on QSO numbers near to galaxy groups and any implications for the total mass present in groups of galaxies, characterised by $\Omega_m$.

A major limitation of the groups of galaxies derived in Chapter 3 is that they are purely selection on the basis of the angular distribution of galaxies on the plane of the sky. In Chapter 4, we use simulated galaxy catalogues to test how many groups selected on the plane of the sky are genuinely groups in three-dimensions and how many are merely

chance alignments on the plane of the sky of distantly-separated galaxies. Recently, Eke et al. (2003) have identified groups of galaxies in the 2dF Galaxy Redshift Survey (2dFGRS) by comparing the distribution of galaxies in the survey to distributions that trace dark-matter haloes in ΛCDM simulations of our Universe. By cross-correlating 2QZ QSOs and 2dFGRS groups, we compare the expected mass of 2dFGRS haloes in ΛCDM simulations to the lensing mass, which we measure under the assumption that QSO numbers are altered in the region of 2dFGRS groups by magnification bias.

As criticisms might still remain about how to define galaxy groups that are truly representative of dark-matter haloes in our Universe, in Chapter 5 we consider a formalism that describes the cross-correlation of QSOs and individual galaxies, rather than groups. The relative number of galaxies around other galaxies is a trace of the *light* in the Universe but (as lensing by dark matter around galaxies influences QSO numbers) the relative number of QSOs around galaxies traces the underyling *mass* in galaxies. Hence a comparison of the galaxy-galaxy auto-correlation to the QSO-galaxy cross-correlation is a comparison of the clustering of light to the clustering of mass. In Chapter 5 we use the QSO-galaxy and galaxy-galaxy correlation functions to trace how biased galaxies are relative to the underlying mass distribution and place constraints on the bias parameter, $b$, on small scales.

Much of the work in this thesis depends on the exact form of the integrated QSO number-magnitude counts fainter than the limit of the 2dF QSO Redshift Survey, which, of course, is a statistic that can't be probed by the 2QZ itself. In Chapter 6, we describe two surveys we have carried out to study the form of the QSO number-magnitude counts in the range $21 \leq b_J < 22$. The 2QZ is now being formally extended to fainter magnitudes and our data from these surveys are a useful test of algorithms used to identify QSO candidates in this fainter survey.

Finally, in Chapter 7 we outline the main results of this thesis and discuss ongoing and future work prompted by our results.

# Chapter 2

# *The 2dF QSO Redshift Survey*

## 2.1 Introduction

The pioneering surveys of QSOs at optical wavelengths (Braccesi, Formiggini & Gandolfi 1970) and the first spectroscopically confirmed QSO samples (Schmidt 1983) provided of the order of 100 QSOs for statistical analyses. As recently as a decade ago, the largest single spectroscopic survey of QSOs (Boyle et al. 1990) identified about 400 QSOs targeted in a sample of $\sim$ 1400 blue star-like objects. The 2dF QSO Redshift Survey represents a phenomenal leap forward in observational technique, largely fuelled by the 2dF spectrograph on the 3.9m Anglo-Australian Telescope (Bailey & Glazebrook 1999, Lewis et al. 2002), which can simultaneously take spectra of 400 objects down optical fibres over a 2° field and uses dedicated software to rapidly reduce the data. Essentially, the 2dF Spectrograph could reproduce the survey of Boyle et al. (1990) in a single night. In full, the 2dF QSO Redshift Survey (Croom et al. 2003) contains 22,159 QSO spectra of high quality (and a further 1181 probable QSOs) observed from a targeted sample of nearly 48,000 objects.

The 2dF QSO Redshift Survey (henceforth 2QZ) area is contained within the 2dF Galaxy Redshift Survey (Colless et al. 2003) and 2QZ QSO candidates were simultaneously observed with 2dF Galaxy Redshift Survey candidates, roughly 275 fibres being reserved for galaxy candidates and 125 for QSO candidates. The 2QZ area comprises 75° in ascension and 5° in declination in both the North and South Galactic Caps (NGC and SGC). In the north, the strip is bounded by $-2.^{\circ}5$ and $2.^{\circ}5$ Declination, and $9^{h}\ 50^{m}$ and $14^{h}\ 50^{m}$ Right Ascension. The southern strip extends from $-32.^{\circ}5$ to $-27.^{\circ}5$ Declination and from $21^{h}\ 40^{m}$ to $3^{h}\ 15^{m}$ Right Ascension. The total area observed is about 740 square degrees. Note that some areas within the boundaries of the 2QZ strips are not surveyed, as objects cannot be identified in the region of bright stars.

QSOs are often characterised by broad spectral emission lines in Lyman-$\alpha$ (1216Å), CIV (1549Å), CIII (1909Å) and MgII (2799Å), and UV continuum emission. Hence, QSOs are often considerably bluer than the majority of stars. Even QSOs at quite large redshifts ($z \lesssim 2.2$) have emission at the blue end of the optical spectrum. The large distance of QSOs means that they appear star-like, so, QSOs can be targeted as objects that are stellar in appearance and have a UV-excess (UVX), provided the QSOs in question are at large enough redshift ($z \gtrsim 0.3$) that they are not reddened or extended by light from their host galaxy. The 2QZ QSO target sample (Smith et al. 1997), was UVX-selected from colour cuts in the $(u - b_J)$, $(b_J - r)$ plane of APM scanned UKST data to limiting magnitude $18.25 \leq b_J < 20.85$. In this thesis, we generally refer to the 2QZ target sample as the "input catalogue" and objects that are present in the input catalogue as "QSO candidates".

Spectra of each object in the 2QZ input catalogue were taken with the 2dF spectrograph, to identify QSOs and separate them from other objects that emit in the UV, mainly, bluer stars, White Dwarfs and some Active Galactic Nuclei, such as Narrow Emission Line Galaxies (NELGs). At the bright end, the survey has been independently extended to $16 \leq b_J < 18.25$ using the 6dF spectrograph on the AAT (we shall refer to this part of the survey as the "6QZ"). 2QZ spectroscopic observations were carried out between 1997 and 2001. Each 2QZ field was observed for around an hour, with many fields overlapping previously observed regions to improve the completeness of the survey. Notably, the minimum fibre separation allowed by the 2dF instrument is of the order of 30 arcseconds, meaning that, had the observational credo not allowed for overlapping fields, no QSOs could have been observed within about 30 arcseconds of a counterpart. The chosen diffraction grating allowed 2QZ spectra to be observed in the wavelength range 3700-8000Å, meaning the broad MgII emission line would be observable to a redshift of $z > 0.3$ and, to help identify fainter QSOs, two broad emission lines (Lyman-$\alpha$, CIII) would be theoretically simultaneously observable to $z < 4.2$. Practically, though, the onset of sky emission at 5577Å(OI) probably limits identification to $z < 2.6$, unless the Lyman-$\alpha$ forest of absorption lines is identified bluewards of the Lyman-$\alpha$ emission line. However, as the forest enters the $u$-band, QSOs become redder in $u - b_j$ and the UVX selection method itself breaks down. The 2QZ has been found to fairly sample the redshift range $0.3 < z < 2.2$ (Outram et al. 2003).

The 2QZ observing process yielded an average signal-to-noise ratio of around 5 in the central wavelength range of the faintest sources, allowing categorical spectroscopic

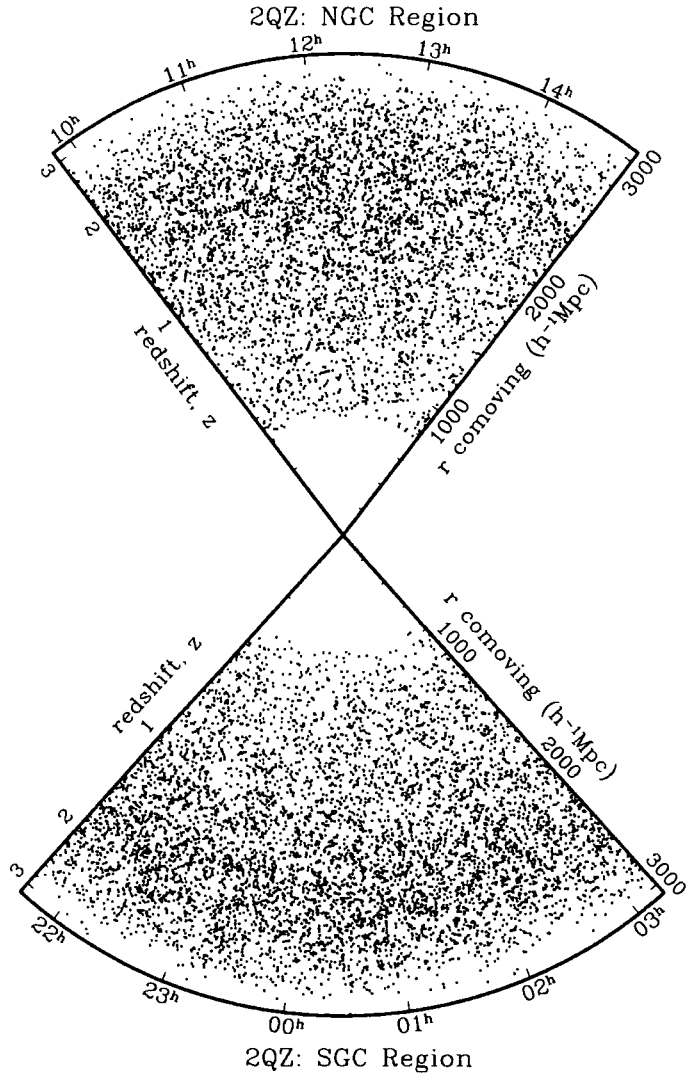


Figure 2.1: The 2dF QSO Redshift Survey (2QZ) identified around 22,000 QSOs in two strips of sky, one in the South Galactic Cap (SGC) and one in the North Galactic Cap (NGC). QSOs in these two regions are plotted by their Right Ascension (B1950) along the curved axis and their redshift, $z$ along the axis running bottom left to top right. A third axis shows the comoving distance of the QSOs assuming an Einstein-de Sitter cosmology. *The software used to make this plot is based on a macro provided by Scott Croom of the Anglo-Australian Observatory.*

identification of 22,159 QSOs, 9856 stars, 4453 NELGs, 2029 White Dwarfs and about 100 other galaxies and AGN. Of the QSOs, 12,303 were identified in the SGC strip and 9856 in the NGC strip. The distribution of 2QZ QSOs by redshift and Right Ascension is plotted in Fig. 2.1. For further general information on the technicalities of the survey, consider Smith et al. (1997), Boyle et al. (2000) or Croom et al. (2001).

In this Chapter, we discuss elements of the 2QZ that will have most bearing on our work in this thesis. In Section 2.2 we discuss the construction of masks that can be used to recreate the angular distribution of QSOs in the survey, were the QSOs free from any clustering effects. In Section 2.3 we derive results from the 2QZ survey that will be pertinent in this thesis, including the number-magnitude and number-redshift relationships for 2QZ QSOs. Finally in Section 2.4 we summarise the important aspects of this chapter.

## 2.2 The Angular Selection Function of the 2dF QSO Redshift Survey.

Much of this thesis will study the relative positions of galaxies and QSOs. To make meaningful statements about the relative clustering of QSOs, it is useful to compare the QSO distribution to a set of points in space that are randomly distributed but with the same positional constraints imposed upon them as 2QZ sources - a "random catalogue". The most obvious example of a random QSO catalogue would be a set of points that were randomly distributed within the boundaries of the 2QZ NGC and SGC strips (except where cuts made around bright stars) - as there can be no 2QZ QSOs in regions outside the survey boundaries (or in regions excised from the survey). In reality, the situation is more complex. Consider a survey that consists of two fields. One field is observed once for an hour. The second field is observed twice, for a total of two hours. If we wish to create a random catalogue that adheres to the same observational constraints as QSOs observed by the 2QZ, shouldn't we populate the second field with more random QSOs than the first? If hundreds of overlapping fields are observed, how do we characterise the likelihood of a QSO in any region of the 2QZ being surveyed? What is the "angular selection function" of objects in the 2QZ?

The completeness of any field in the 2QZ is characterised as a product of two other types of completeness:

- *Observational completeness* - A field contains 200 QSO candidates but only 100 have been observed by a 2dF fibre. The *observational completeness* of the field is then 100/200.

- *Spectroscopic Completeness* - Of the 100 QSO candidates that have been observed by a 2dF fibre, 75 have reliable spectra - perhaps 15 identified as stars, 45 as QSOs, 10 as NELGs and 5 as White Dwarfs. The other 25 spectra are of poor quality and have no observation. The *spectroscopic completeness* of the field is then 75/100.

- *Completeness* - So, in total, 75 objects have been identified from 200 candidates. The *completeness* of the field is 75/200.

Again, the situation is slightly more complicated than outlined above for the following two reasons

- The quality of spectra are not necessarily simply "reliable" or "poor". One observer may find absorption lines in a spectrum that a second observer misses. All spectra in the 2QZ have been identified independently by two different observers. If an observer considers a spectrum to be of good quality, they assign it a quality of "1". If they consider it of poor quality, they assign it a quality of "2". Objects with no identification are assigned a quality of "3". 2QZ objects, then, have a spectral quality consisting of a two-digit number, such as "12", which would mean that the first observer thought it of good quality and the second observer thought it of poor quality. In Section 2.1, when we listed numbers of objects in the 2QZ with "reliable identifications", we were referring to objects of "11" quality. Note, that the spectroscopic completeness of the 2QZ is hence dependent upon a chosen cut in spectral quality. It might be the case that 75/200 objects observed in a given field are of "22" spectroscopic quality (or better) but only 60/200 objects are of "11" quality.

- The calculation of spectroscopic completeness depends on the assumption that the relative fractions of objects with no spectra are fair reflections of the relative fractions of objects that have been identified. *A priori*, it might seem likely that QSOs and NELGs, which have obvious emission lines, might be readily identified and most unidentified objects would, therefore, be stars. To test this supposition, we have considered every object that was assigned a "33" spectroscopic quality when first observed as part of the 2QZ but was subsequently observed a second time and

| | Entire 2QZ (%) | | $b_J > 20.5$ (%) | |
|---|---|---|---|---|
| ID | All IDs | Only re-IDs | All IDs | Only re-IDs |
| QSO | $58.99 \pm 0.40$ | $58.73 \pm 1.46$ | $47.20 \pm 0.65$ | $50.09 \pm 2.07$ |
| NELG | $11.85 \pm 0.18$ | $7.63 \pm 0.52$ | $23.03 \pm 0.45$ | $12.61 \pm 1.04$ |
| star | $23.50 \pm 0.25$ | $28.80 \pm 1.02$ | $26.34 \pm 0.48$ | $33.56 \pm 1.69$ |
| WD | $5.40 \pm 0.12$ | $4.41 \pm 0.40$ | $3.00 \pm 0.16$ | $2.98 \pm 0.50$ |
| others | $0.26 \pm 0.03$ | $0.43 \pm 0.13$ | $0.43 \pm 0.067$ | $0.77 \pm 0.26$ |

Table 2.1: 2767 objects surveyed in the 2QZ were initially assigned a "33" spectroscopic quality after their first observation (meaning they could not be identified), then were subsequently reobserved and reidentified with a "11" spectroscopic quality. In columns 2 and 3, we display fractions of identifications in the 2QZ for the entire survey and for the 2767 reidentifications. In columns 4 and 5, we repeat this analysis for faint objects in the 2QZ ($b_J > 20.5$). Even for $b_J > 20.5$, we cannot reject the hypothesis that the fraction of QSOs in the 2QZ as a whole is reflective of the reidentified sample. Errors in this table are Poisson.

reidentified with a "11" spectroscopic quality. In Table 2.1 we list the fraction of objects in the 2QZ by their identification compared to the fraction for reidentified objects. Even at faint magnitudes ($b_J > 20.5$), where the signal-to-noise is lowest, the fraction of reidentified QSOs is in reasonable agreement with the fraction of QSOs in the 2QZ as a whole. We are thus confident that correcting QSO numbers for spectroscopic incompleteness in the manner we have outlined fairly represents the true numbers of QSOs.

- It is not really practical to calculate survey completeness on the basis of 2QZ fields. As many fields overlap, some areas of the 2QZ might be observed three or four times more often than their counterparts. This is illustrated in Fig 2.2, where three hypothetical 2QZ fields of different individual completeness overlap to form seven sectors. The three points itemised above to outline the calculation of field completeness are actually used to calculate the completeness of each *sector*. We refer to the resulting distribution of the completeness of the 2QZ by position as the angular selection function or, simply, as the 2QZ mask.

So, to create a random catalogue of points with the same angular selection function as the 2QZ, we place random points in 2QZ sectors with the same relative frequency as the relative completeness of the sectors. Areas outside of the 2QZ boundary, or in holes excised for bright stars, have zero completeness. Note the exact form of the mask depends

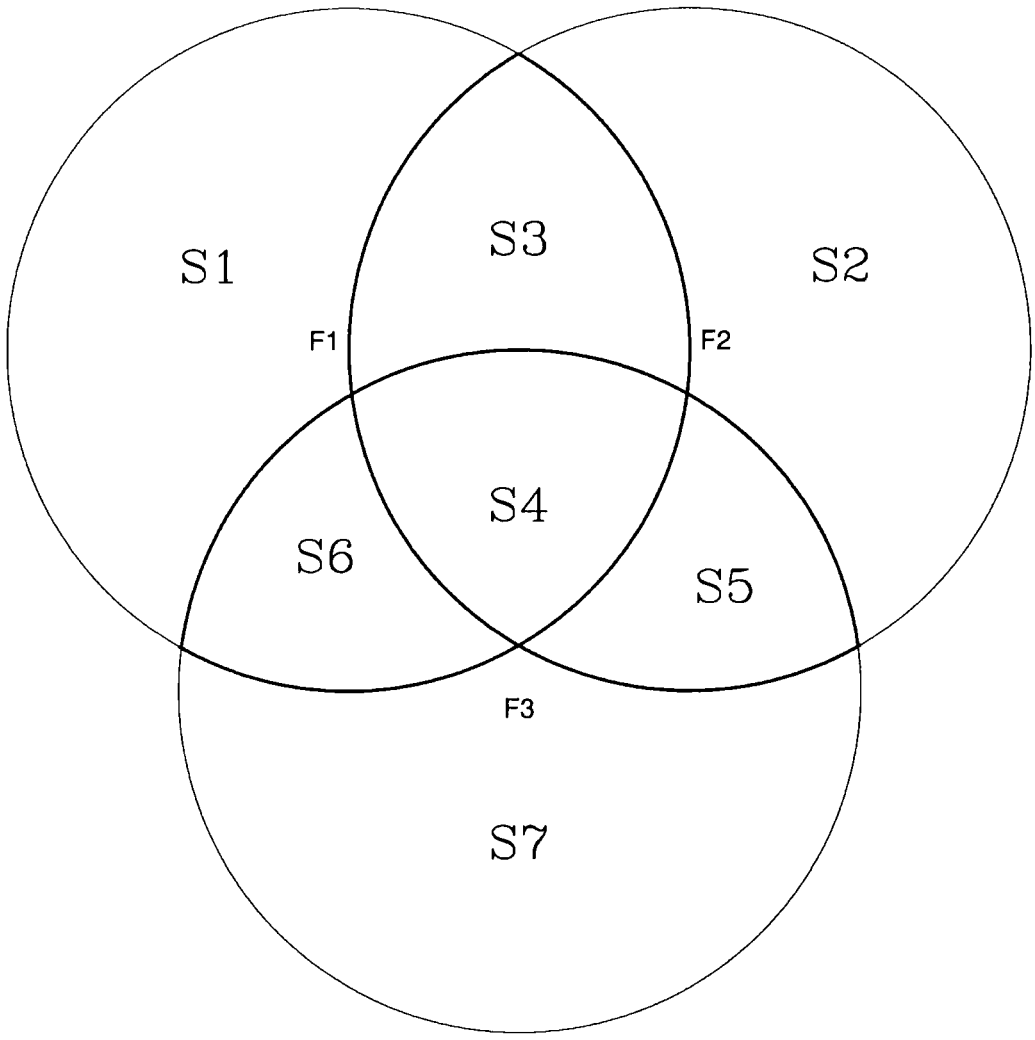


Figure 2.2: A diagram illustrating the schema by which 2QZ fields are divided into sectors to determine the completeness of any area on the sky. The circles F1, F2 and F3 are hypothetical 2dF *fields*, observed on different nights and, therefore, likely of different completeness. The fields divide up into S1, S2, S3, S4, S65, S6 and S7, which would be considered different 2dF *sectors*. The completeness of the 2dF would be calculated in each *sector*, as outlined in the text.

on the nature of the data sample, including its spectroscopic quality and magnitude limit. When considering the clustering of 2QZ QSOs in this thesis, we switch between using "22" quality data and "11" quality data, and consider QSO samples to different magnitude limits. Though we might not specifically state as much, in each case we have recalculated the 2QZ mask before creating a random catalogue of points to compare against the 2QZ QSO distribution. We make masks using software based on a system of mapping 2QZ sectors invented by Dr. Scott Croom of the Anglo-Australian Observatory

Another effect that can influence whether more objects are included in given sectors of the 2QZ is obscuration by dust in our Galaxy. Two 2dF sectors may appear to have identical completeness but, in fact, light from QSOs in one of the sectors is almost entirely obscured by dust. Consequently, the identical completeness indicates that there are actually relatively many QSOs in the sector obscured by dust. Across the NGC and SGC strips in their entirety, dust is not really an issue. The strips are chosen to lie in areas of low Galactic absorption. According to the maps of Schlegel, Finkbeiner & Davis (1998), the mean absorption in $b_J$ across the SGC strip is 0.077 mags. Absorption is slightly higher in the NGC strip (a mean of 0.145 mags in $b_J$). However, fluctuations in the amount of dust can be a problem when comparing the clustering in two independent areas of the 2QZ. When making a random catalogue we can correct for dust at every point using the maps of Schlegel, Finkbeiner & Davis (1998) and weighting the likelihood,$W$, of a random point (of Right Ascension $\alpha$ and Declination $\delta$), by dust absorption $A_{b_J}$ along that line of sight, according to

$$W(\alpha, \delta) = 10^{-\alpha A_{b_J}} \tag{2.1}$$

where $\alpha$ is the value of the integrated slope of the QSO number-magnitude counts at the magnitude limit of the 2QZ. We discuss the value of $\alpha$ in Section 2.3.

In Figure 2.3 we illustrate some of the points made in this section by plotting a dust map, and 2QZ and 6QZ completeness masks for areas of the 2QZ NGC strip. In both cases, the completeness masks are calculated for objects with a "11" spectroscopic quality or better and to the magnitude limit of the surveys ($b_J < 18.25$ for the 6QZ and $b_J < 20.85$ for the 2QZ).

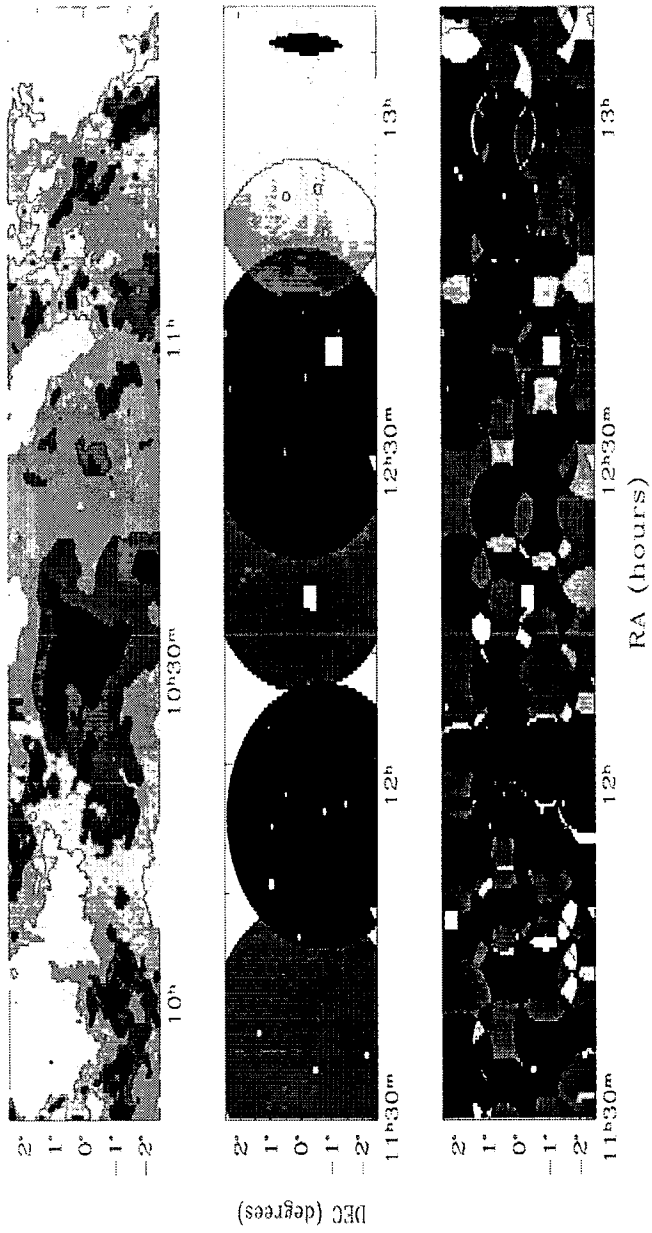


Figure 2.3: The upper panel (when viewed landscape) shows dust in part of the 2QZ NGC strip (Schlegel, Finkbeiner & Davis 1998), binned by the amount of absorption in $b_J$ magnitudes. Darker shades represent more absorption - black for absorption of $> 0.25$, white for $< 0.1$, with three bins of 0.05 in between. The middle panel shows a region of the 6QZ by completeness. Darker shades in this panel represent higher completeness - black representing $> 80$ per cent of objects were observed, white for $< 20$ per cent observed, with three bins of 20 per cent in between. The lower panel (when viewed landscape) is the same as the middle panel but for the 2QZ.

## 2.3 Number Counts By Magnitude and Redshift

As outlined in Appendix B, the expected strength of lensing-induced correlations between galaxies and a magnitude-limited sample of QSOs depends on the slope of the integrated number-magnitude counts, $\alpha$, fainter than the QSO sample's limit (Narayan 1989). We need to estimate this slope to interpret some of the results recorded in this thesis. As we do not have information fainter than the limit of the 2QZ, we fit models to the 2QZ number-counts brighter than the $b_J = 20.85$ limit and extrapolate the counts to fainter magnitudes. Such models are better constrained by fitting them to the differential counts and integrating them, as the differential counts are well known across the range of the 2QZ and 6QZ, while the integrated counts in brighter bins also depend slightly on QSO numbers brighter than the 6QZ. However, as practically we are interested in the faint-end slope of the integrated counts, we fit the model that best approximates both the integrated and differential number-counts.

In Fig. 2.4 we present the differential number-counts by magnitude of the 2QZ. The counts have been corrected for incompleteness and absorption by dust in our Galaxy, and averaged over both hemispheres. Also plotted are points from the 6QZ. When determining the number-count relation, the sample is restricted to the redshift range $0.3 < z < 2.2$, the range for which the 2QZ is designed to be photometrically complete (Croom et al. 2001) and appears to be homogenous (Outram et al. 2003).

The dashed line is a Smoothed Power Law (SPL) model, where the differential counts are expressed in the form

$$\frac{\mathrm{d}N}{\mathrm{d}m} = \frac{N_0}{10^{-\beta_\mathrm{d}(m-m_0)} + 10^{-\alpha_\mathrm{d}(m-m_0)}} \tag{2.2}$$

The best-fit model has a bright-end slope of $\beta_\mathrm{d} = 0.98$, a knee at $m_0 = 19.1$ and a faint-end slope of $\alpha_\mathrm{d} = 0.15$. This model is consistent with faint data from Boyle, Jones & Shanks (1991) and Koo & Kron (1988), which are also marked in Fig. 2.4. Data from the literature have undergone a zeropoint correction of $b_J = B - 0.1$ and a dust correction of about 0.1 magnitudes (the average absorption in the regions surveyed). Note that we do not use these data fainter than $b_J = 20.85$ in the best-fit SPL model, as they have negligible statistical power compared to the 2QZ and 6QZ.

In Fig. 2.5, the SPL model is integrated and displayed against the integrated QSO number-magnitude counts. The best-fit SPL model has an average integrated faint-end slope (over the range 1 magnitude fainter than the limit of the 2QZ) of $\alpha = 0.29$. Although

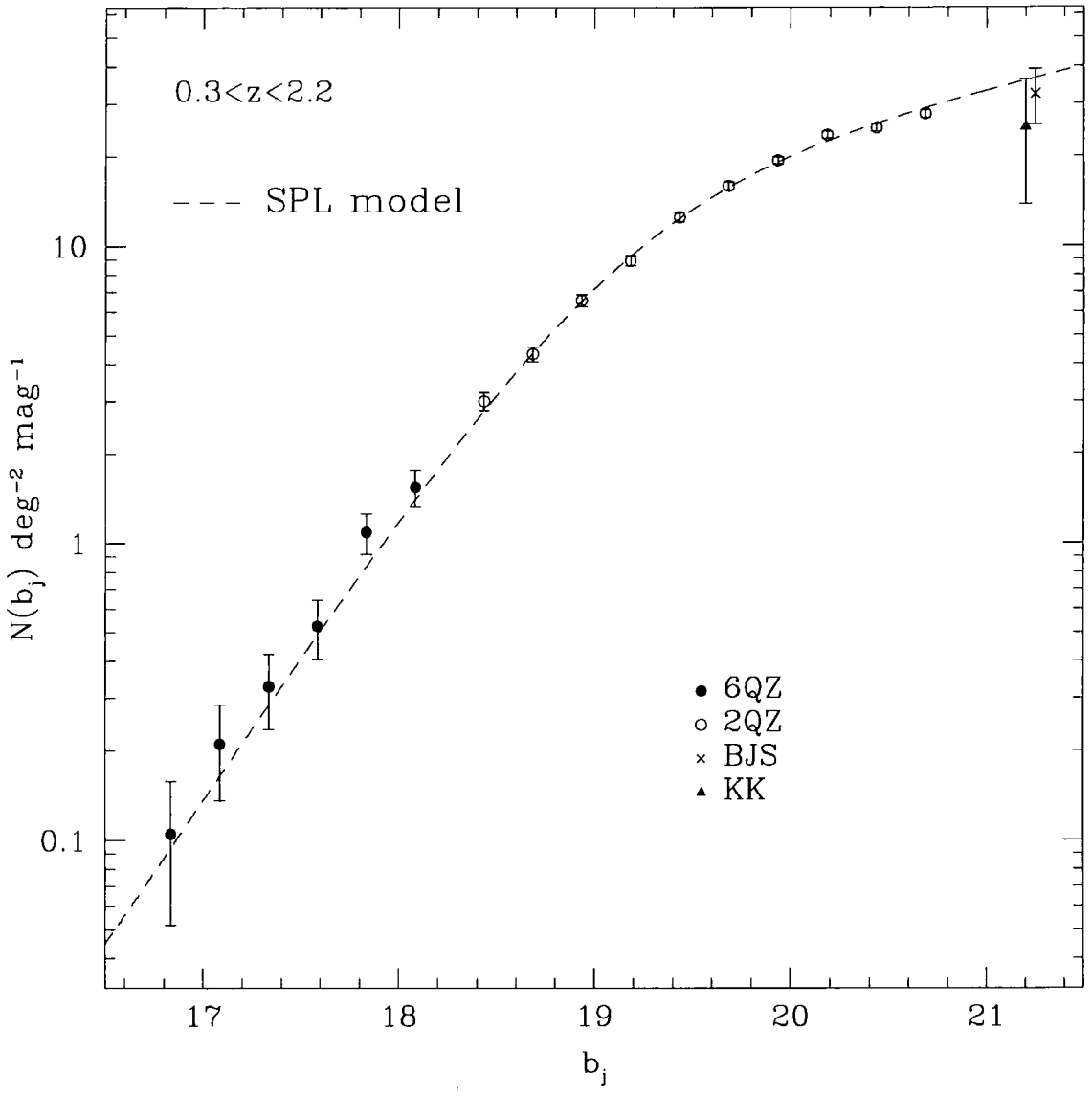


Figure 2.4: The differential number-counts of QSOs identified in the 2QZ. The points are QSO number-counts in 0.2 mag bins, with Poisson errors. The dashed line is a smoothed power law fit that is the best compromise fit to both the differential and integrated QSO number-counts. Brighter data points are from the 6QZ. Also displayed are the faintest data from Boyle, Jones & Shanks (1991) and Koo & Kron (1988).

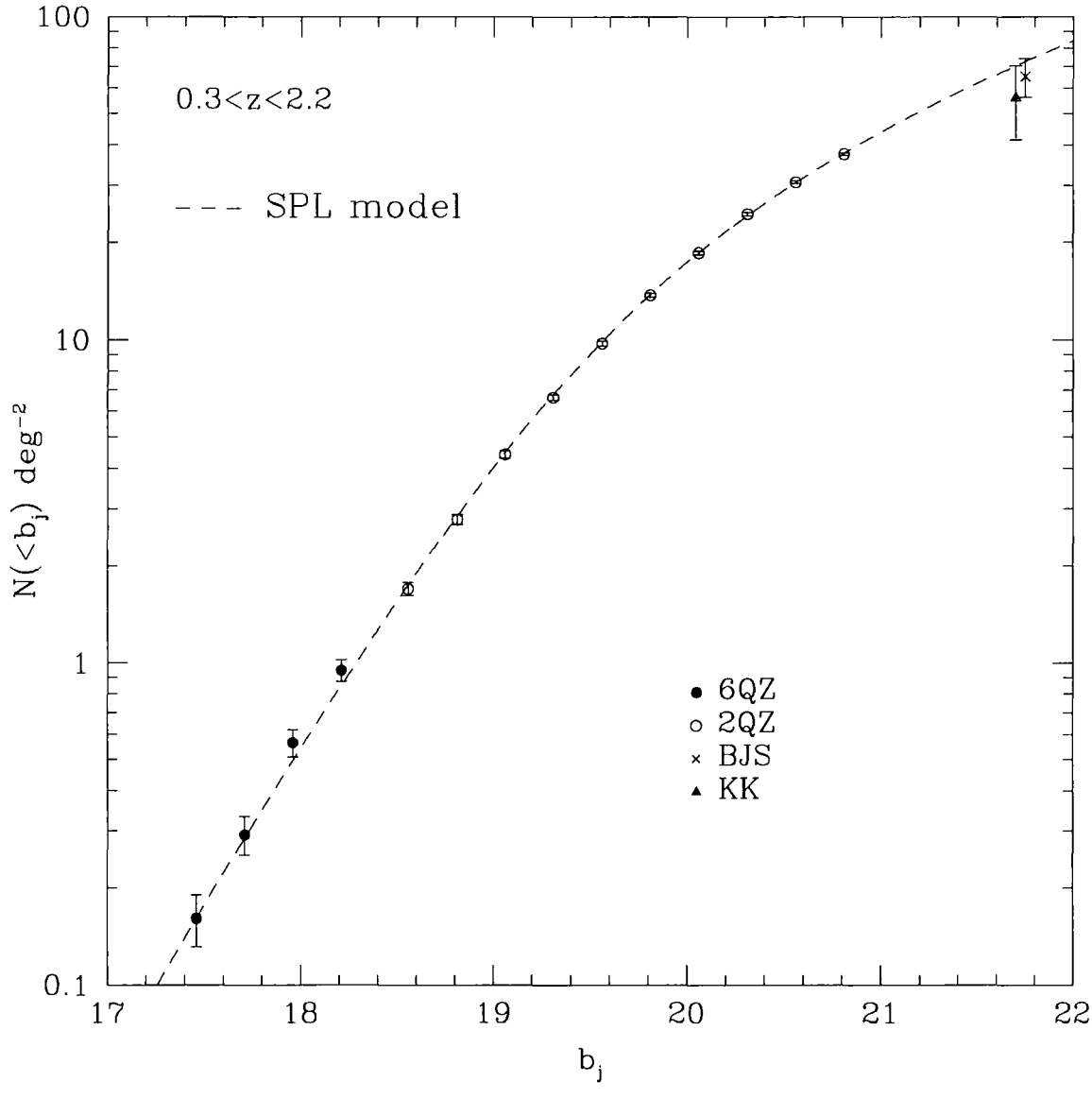


Figure 2.5: The $N(< m)$ relation, or integrated number-counts, for the 2dF QSO Redshift Survey. The points are QSO number-counts in 0.2 mag bins, with Poisson errors. The line is derived by integrating the smoothed power law model for the differential number-counts and represents the best compromise fit to both the integrated and differential QSO number-counts. Brighter data points are from the 6QZ. Also displayed are the faintest data from Boyle, Jones & Shanks (1991) and Koo & Kron (1988), which have been offset slightly to prevent the points from merging.

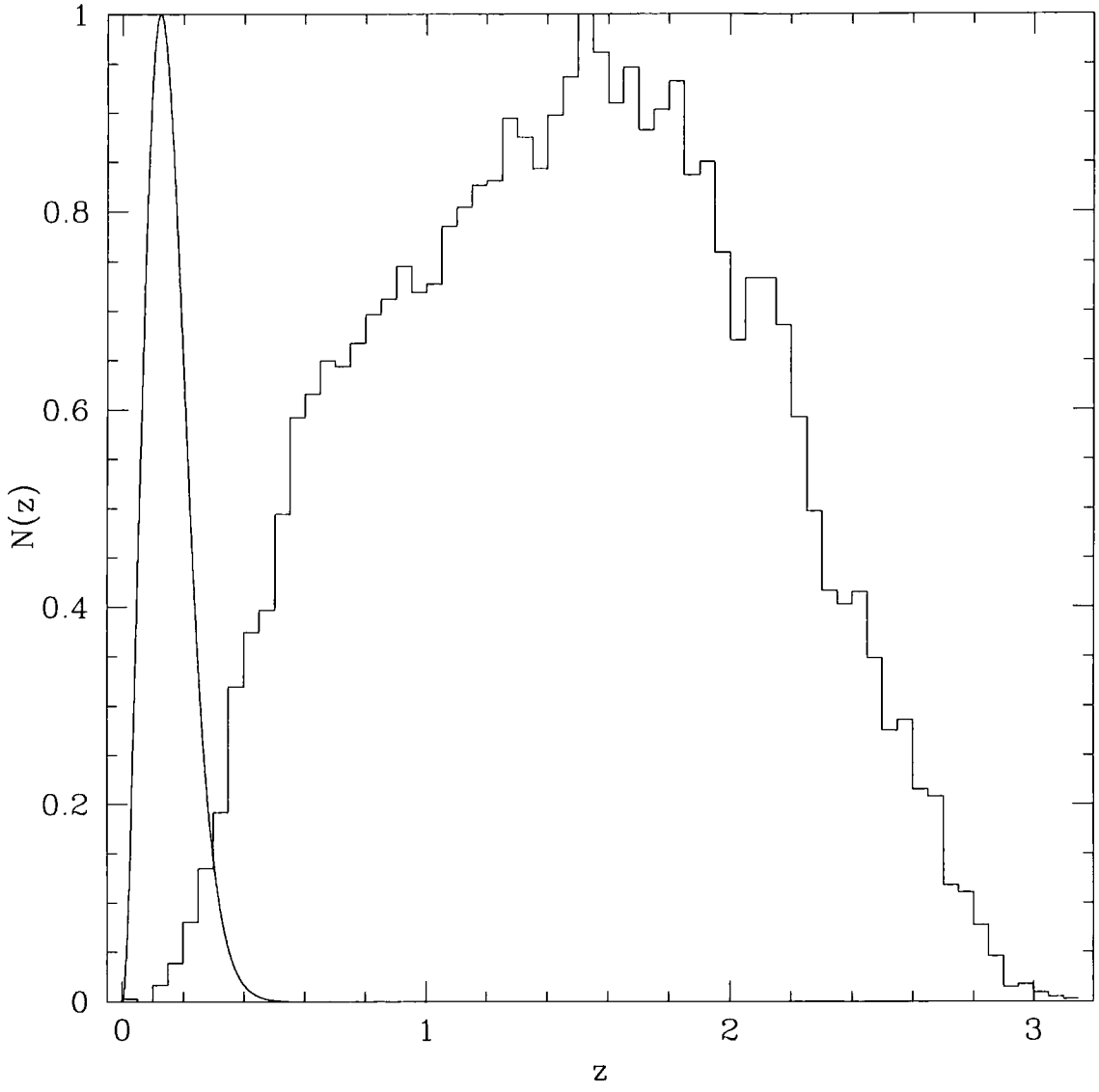


Figure 2.6: The N(z) relationship for the 2dF QSO Redshift Survey. The wider histogram represents QSO number-counts in 0.05 redshift bins. The narrow histogram is an analytic model of the galaxy distribution at the $b_J = 20.5$ limit of the galaxy samples typically used in this thesis. The distributions are normalised to peak at unity.

the model is well constrained by the data ($\pm 0.02$) the many incompleteness corrections to the faint-end data mean the $1\sigma$ error may be as high as $\pm 0.05$. When Boyle, Fong & Shanks (1988) modelled the magnitude distribution of QSOs as a broken power law, they determined a B-band faint end slope of $0.32 - 0.33$. In an extensive review, Hartwick & Schade (1990) subsequently determined a faint-end slope of 0.31. Our average slope is thus consistent with these earlier authors.

When making model predictions to interpret results in this thesis, we shall also need to know the redshift distributions of both QSOs and galaxies. In Fig. 2.6 we display the number-counts by redshift in the 2QZ as a histogram. Also marked is an analytic model for a galaxy redshift distribution (Baugh & Efstathiou 1993) integrated to $b_J = 20.5$, the galaxy sample limit we typically consider in this thesis. Both distributions have been normalised to peak at 1. Note that for redshifts greater than 0.4, less than 0.4 per cent of the projected galaxy distribution overlaps the QSO distribution.

## 2.4 Summary

We have introduced the 2dF QSO Redshift Survey (2QZ) and 6dF QSO Redshift Survey (6QZ). We have discussed aspects of 2QZ data, notably the spectroscopic quality of data, which can take a value of "11", "12", "22", "23" or "33" depending on a judgement of quality in the range 1-3 made by two different observers. We have discussed the partitioning of the 2QZ survey into sectors and the calculation of the observational, spectroscopic and total completeness of a given sector. We have noted that the fraction of QSOs that were observed but not not identified in the 2QZ is likely to be a fair reflection of the fraction of QSOs that were identified in the 2QZ, meaning that QSO numbers can be fairly corrected using an estimate of spectroscopic completeness based on the ratio of identified to unidentified objects in a 2QZ sector. We displayed examples of maps of the completeness of the 2QZ as a function of postion - the angular selection function of the 2QZ - which can be used to make a random catalogue of 2QZ objects. We have also discussed the influence of dust on the angular completeness of the 2QZ. Finally, we have displayed the QSO number-magnitude and number-redshift counts and derived the faint-end slope of the integrated QSO number-magnitude counts. In subsequent chapters, aspects of the 2QZ discussed in this chapter will be used to study lensing of QSOs by foreground structure.

# Chapter 3

# *The Correlation of Faint QSOs and Galaxy Groups*

## 3.1 Introduction

Connections between low-redshift structure and high-redshift QSOs were the subject of some debate for decades (Burbidge 1979). Chance associations are statistically unlikely (Kippenhahn & de Vries 1974, Burbidge, Hoyle & Schneider 1997), prompting some authors to theorise the proximity of high-redshift QSOs to low-redshift galaxies in keeping with a non-cosmological interpretation of QSO redshifts (Burbidge et al. 1990, Burbidge & Hoyle 1996). Such a hypothesis has become increasingly unlikely for a variety of reasons, not least the continuing correlation between high redshift Damped Lyman Alpha systems (DLAs) and galaxies; see Ellison et al. 2001) and alternative explanations have become increasingly popular.

Claims of associations between QSOs and foreground structure are now more usually interpreted in terms of gravitational lensing (Canizares 1981) than evidence for non-cosmological redshifts (Bukhmastova 2001, Benítez, Sanz & Martínez-González 2001). Observationally, however, the situation has been complex, with a variety of effects recorded in the literature. For example Williams & Irwin (1998) detected a positive correlation between high-redshift LBQS QSOs and APM galaxies on degree scales, yet Martinez et al. (1999) find no strong quasar-galaxy angular correlation on similar scales. Ferreras et al. (1997) found a large anti-correlation between galaxies and optically selected QSOs near the NGP, and suggested a difficulty in selecting QSOs in densely populated areas. Samples of QSOs utilised in these and earlier papers frequently suffered from either inhomogeneity or a dearth of data (see, e.g., Norman & Williams 2000, for a review). The completed 2dF QSO Redshift Survey contains a UVX-selected homogeneous sample of

around 23,000 QSOs. Contemporary models suggest that statistical lensing should become a stark cosmological effect in such vast surveys (Menard & Bartelmann 2002).

Following Shanks et al. (1983), Boyle, Fong & Shanks (1988) quoted a significant anti-correlation between objectively-selected groups of galaxies and faint UVX objects. Initially, Boyle, Fong & Shanks (1988) interpreted the anti-correlation as an effect caused by a small amount of dust in foreground galaxy groups obscuring background QSOs. Ferguson (1993) and Maoz (1995) generally restricted reddening in clusters and rich groups, results that marginally suggested insufficient dust to induce the observed lack of QSOs around galaxy groups. This prompted Croom & Shanks (1999) to recast the anti-correlation signal in terms of statistical lensing. Rodrigues-Williams & Hogan (1994) first discussed the Boyle, Fong & Shanks (1988) anti-correlation result as a possible effect of gravitational lensing in a paper confirming that bright QSOs are positively correlated with galaxy clusters. Significant positive correlations between galaxies and bright QSOs, which cannot be explained away by dust in galaxies, continue to be detected (Williams & Irwin 1998, Norman & Williams 2000, Norman & Impey 2001, Gaztañaga 2003).

Gravitational lensing can satisfactorily explain associations between QSOs and foreground structure. As discussed in Appendix C, interim mass lenses the area behind it, influencing a sample of distant objects in two related ways. Firstly, sources are magnified. Secondly, the apparent sky density of sources of given intrinsic luminosity drops. If the number-magnitude relation of the considered sample is steep, the first effect dominates and we observe the more numerous, fainter QSO population. If the number-magnitude slope is shallow, the second effect prevails and we see fewer QSOs, as the area that they populate dilutes. Ultimately, this magnification bias will increase the correlation of QSOs and galaxies if we consider a sample near an intrinsically steep part of the QSO number-magnitude relation, and it will induce a paucity of QSOs around foreground mass where QSO number-counts flatten (Kovner 1989, Narayan 1989, Schneider 1989, Wu 1994). The optical integrated QSO number-counts are steep for bright QSOs and flatten significantly at faint magnitudes (see Fig. 2.5). Observationally, then, we would expect both positive and negative correlations between QSOs and galaxies, depending on the apparent luminosity of the QSO sample.

Questions remain about the anti-correlation between faint UVX objects and galaxies in groups detected by Boyle, Fong & Shanks (1988). Is the result reproduced in a different sample? Is it affected by dust in galaxy groups? Is it a selection effect or a systematic? Is statistical lensing a viable explanation given the large amplitude of the anti-correlation?

In this paper, we address these questions by measuring the two-point correlation function between objects in the 2dF QSO Redshift Survey (Croom et al. 2001) and galaxy groups determined from the APM catalogue (Maddox et al. 1990a) and from the Sloan Digital Sky Survey Early Data Release (Stoughton et al. 2002). As discussed in Chapter 2, the 2dF QSO Redshift Survey (2QZ) contains a large, homogeneous, sample of objects. The QSOs are spectroscopically confirmed, meaning contamination is extremely low. The 2QZ also measures colour, allowing limits to be placed on the effect of dust in galaxies on background QSOs. Finally, in modelling our lensing signal, we consider a different cluster profile to Boyle, Fong & Shanks (1988) and adapt our analysis so that the model more fairly represents the data.

In this chapter, we investigate the cross-correlation between QSOs and foreground galaxy groups. In Section 3.2 we outline the galaxy samples from which we define groups of galaxies to cross-correlate with QSOs and the objective method by which we derive galaxy groups. In Section 3.3 we present our cross-correlation analysis, discussing the possibilities that its form is attributable to either selection effects or dust. In Section 3.4 we introduce a model of statistical lensing of background QSO light by foreground dark matter haloes, which could lead to a cross-correlation arising between QSOs and galaxy groups. In Section 3.5 we interpret our results as indicative of statistical gravitational lensing and discuss the implications of such an interpretation. Section 3.6 presents our conclusions.

## 3.2 APM and SDSS clusters

### 3.2.1 The catalogues

Our correlation analysis in subsequent sections relies on groups of galaxies drawn from two generations of surveys, the APM Galaxy Survey (Maddox et al. 1990a) and the Early Data Release (Stoughton et al. 2002) of the Sloan Digital Sky Survey (henceforth SDSS EDR).

The initial APM survey (Maddox et al. 1990a) was derived from Automated Plate Measuring scans of 185 photographic UKST plates and covered about 10 per cent of the entire sky around the Southern Galactic Cap (henceforth SGC). The original APM region was bounded between a Right Ascension of roughly $21^{\rm h}$ to $5^{\rm h}$, with Declination from $-72\,^\circ$ to $-18\,^\circ$. Images were detected to $b_J < 21.5$, allowing galaxies to be identified to $b_J < 20.5$. The photometry was aligned using overlapping plates (Maddox et al. 1990b).

The APM Survey was later extended and provided the input catalogues for the 2dF Surveys. The original APM Survey completely covers 2QZ SGC strip.

The SDSS (`http://www.sdss.org`) is imaging the northern sky in five bands designed for CCD photometry ($u, g, r, i$ and $z$). The survey will trace an ellipse centred on $12^{\rm h}\ 20^{\rm m}$ Right Ascension and 32.8 ° Declination, roughly extending from $7^{\rm h}\ 6^{\rm m}$ to $17^{\rm h}\ 34^{\rm m}$ in Right Ascension and ±55 ° of declination. The SDSS should be complete to $g' \sim 23.3$ and $r' \sim 23.1$ (York et al. 2000), about equivalent to $b_J \sim 23.5$ using a typical colour transformation (Yasuda et al. 2001). Note that the dashes on the magnitude bands indicate Early Data Release (EDR) photometry, which is preliminary. The EDR (Stoughton et al. 2002) contains the first 460 deg$^2$ of the SDSS. Conveniently, one EDR strip, extending from $9^{\rm h}\ 40^{\rm m}$ to $15^{\rm h}\ 44^{\rm m}$ in Right Ascension and −1 ° to 1.5 ° Declination, crosses the 2QZ NGC strip.

### 3.2.2 Objective group catalogues

We wish to correlate the 2QZ with galaxy groups rather than the field as any correlation attributed to lensing will be stronger for more massive foreground structures. A catalogue of objectively identified groups in the Southern APM is available in the literature (Dalton et al. 1997) but comparatively little has been available in the region of the northern 2QZ since the publication of the ACO catalogue (Abell, Corwin & Olowin 1989). Additionally, it is useful to have listings of the position of each galaxy within each cluster, not merely the positions of cluster centres. Following Boyle, Fong & Shanks (1988) we turn to the clustering routine of Turner & Gott (1976) to objectively identify clusters of galaxies in the APM survey and SDSS EDR.

To determine the boundaries for cluster membership in a galaxy sample, we assign an overdensity, $\delta$, a factor by which we wish our group density to exceed the mean surface density ($\bar{\sigma}$) across the entire region we consider. We then calculate the largest possible circle of radius $\theta_c$ such that

$$\sigma(\theta < \theta_c) > \delta\bar{\sigma} \tag{3.1}$$

where $\sigma$ is the surface density of galaxies within the (circular) region centred on any particular galaxy in our sample and enclosed by an angular radius $\theta$. Over the miniscule angles typically considered, $\sigma$ can be expressed as

$$\sigma = \frac{N(\leq \theta)}{2\pi(1-\cos\theta)} \approx \frac{N(\leq \theta)}{\pi\theta^2} \tag{3.2}$$

where $N(\leq \theta)$ is the total number of galaxies within an angle $\theta$ of the particular galaxy we are considering, including the particular galaxy itself. The critical angular radius $\theta_c$ is determined for each individual galaxy in the sample. A group is defined as all galaxies that have overlapping critical radii. Note that a typical angular radius for any group could be estimated via

$$\theta_t = \sqrt{\frac{N}{\pi\bar{\sigma}\delta}} \tag{3.3}$$

Two classification questions remain; how overdense are groups (what value should $\delta$ take)? and what size must groups attain before we call them a group (what is the minimum number of galaxies, $N_{min}$, in a group)? We take the values of $\delta = 8$ and $N_{min} = 7$ chosen by Stevenson, Fong & Shanks (1988) in a similar analysis, and used by Boyle, Fong & Shanks (1988). The choice of the overdensity parameter was originally suggested by Turner & Gott (1976) and weighs the possibility of losing poor clusters at high $\delta$ values against false grouping of galaxies in the field at low $\delta$. The choice of $N = 7$ as a minimum group size is intended to reduce the likelihood of chance alignments of galaxies at different redshifts being grouped. The parameters reflect a distribution of groups that may have been selected by eye (Stevenson, Fong & Shanks 1988).

Fig. 3.1 and Fig. 3.2 show the groups determined from the SDSS EDR data in the NGC region of the 2QZ and from APM data in the 2QZ SGC strip. In both cases, the galaxy sample was limited to $b_J < 20.5$, the limit of the APM. The SDSS EDR data was transformed from the SDSS photometric system using the colour transformations of Fukugita et al. (1996), a procedure that we will discuss in more depth in Chapter 6. The axes in both figures correspond to the limits of the 2QZ. The SDSS EDR coordinate system has been transformed from J2000 to B1950. In general, unless otherwise specified, all coordinates listed in this thesis will be in the B1950 system. In Fig. 3.1 and Fig. 3.2, we also plot the positions of Abell clusters (Abell 1958, Abell, Corwin & Olowin 1989) in the 2QZ strips. In Fig. 3.2, we also plot clusters defined from APM data by Dalton et al. (1997).

A great deal of the structure produced by Dalton et al. (1997) is reproduced well in our SGC group catalogue, even though Dalton used different clustering criteria. 30 per cent

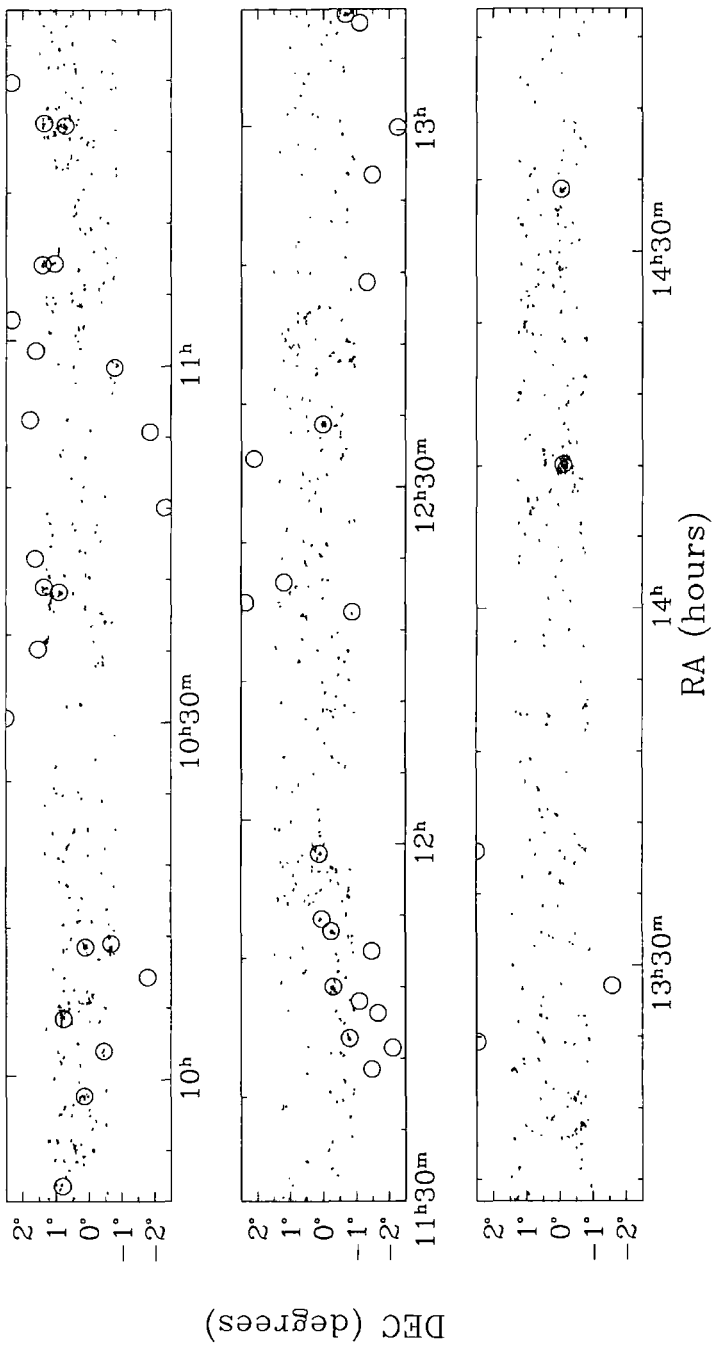


Figure 3.1: Objectively defined galaxy groups in the Sloan Digital Sky Survey Early Data Release. Each point is a galaxy in a region sufficiently dense that it meets clustering criteria outlined in the text. For comparison, Abell clusters (Abell 1958, Abell, Corwin & Olowin 1989) are plotted as circles. The axes correspond to the limits of the 2dF QSO Redshift Survey. The coordinate system is B1950.

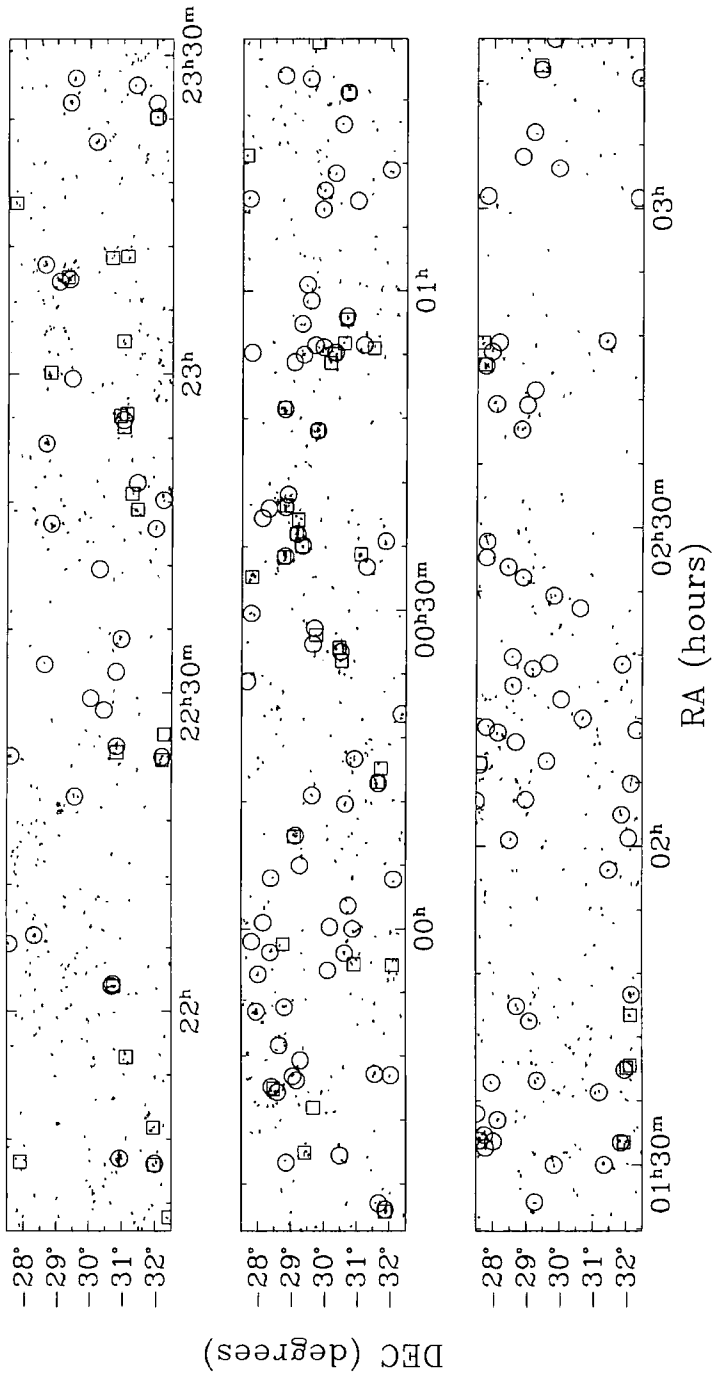


Figure 3.2: Objectively Defined Galaxy Groups in the Southern APM Survey. Each point is a galaxy in a region sufficiently dense that it meets clustering criteria outlined in the text. For comparison, Abell clusters (Abell 1958, Abell, Corwin & Olowin 1989) are plotted as circles and APM clusters (Dalton et al. 1997) are plotted as squares. The axes correspond to the limits of the 2dF QSO Redshift Survey. The coordinate system is B1950.

of Dalton's cluster centres within the 2QZ boundary lie within 2 arcminutes of our group centres, improving to 75 per cent of Dalton's cluster centres lying within 5 arcminutes of our own. Over 90 per cent of group centres match up within a separation of 7.7 arcminutes, the mean cluster radius of Dalton's groups within the 2QZ SGC region. Many of the larger groups automatically picked out by our technique also correspond to previously identified rich clusters (Abell 1958), for instance, the large cluster around $14^h\ 12^m$ Right Ascension in Fig. 3.1 corresponds to Abell 1882 (Richness Class III, $z \sim 0.137$). Abell assigned 166 galaxies to Abell 1882, we determine 153 members. We also pick out voids that are in good agreement with the literature. The sparse region we pick out in the Southern APM from Right Ascension $2^h\ 30^m$ to $3^h$ appears to be real. In this region, Dalton et al. (1997) finds no clusters over the Declination range $-32.^\circ5$ to $-28^\circ$ and Abell finds a single (Richness Class 0) cluster in the range $-32.^\circ5$ to $-29.^\circ5$. The void does not appear to be an effect of intervening dust obscuring galaxies out of the APM - although there is quite high absorption ($> 0.25$ mags) in the most southerly $2^\circ$ of sky betwen Right Ascension $2^h\ 35^m$ and $2^h\ 40^m$, the region of the SGC strip from $23^h$ to $23^h\ 30^m$ generally has both more dust across it (according to the maps of Schlegel, Finkbeiner & Davis 1998) and a denser population of groups. In any case, we note that similar voids are picked out at, say, $22^h\ 30^m$ in the Southern APM, or $10^h\ 30^m$ in the SDSS EDR.

## 3.3 Cross-correlation analysis

We now turn our attention to the cross-correlation of galaxy groups and 2QZ objects. Throughout this chapter, we consider the subsample of 2QZ objects with a spectroscopic quality of "22" or better (2QZ spectroscopic quality is described in Chapter 2). This is unique to this chapter - in subsequent chapters we use "11" spectroscopic quality objects. The clustering results quoted in this chapter are not influenced by this choice of spectroscopic quality (the choice to study "22" quality objects was only made as much initial clustering analysis was carried out well before the completion of the 2QZ, when there were a great deal fewer "11" objects).

### 3.3.1 Method

As discussed in Appendix B, the two-point angular correlation function $\omega(\theta)$ (Peebles 1980) measures the probability ($\mathrm{d}P$) of finding pairs of sources of mean number density $\bar{n}$ within solid angle $\mathrm{d}\Omega$ separated by angle $\theta$

$$dP = \bar{n}_1 \bar{n}_2 (1 + \omega(\theta)) d\Omega_1 d\Omega_2 \tag{3.4}$$

Throughout this thesis, to measure the two-point correlation function, $\omega(\theta)$, we use the estimator derived by Peebles & Hauser (1974) to study the Zwicky galaxy catalogues

$$\omega_{12}(\theta) = \frac{D_1 D_2(\theta) \bar{n}}{D_1 R_2(\theta)} - 1 \tag{3.5}$$

where $D_i$ denotes a data point ($D_1 D_2$ denotes a data-point *pair*) of population $i$ and $R_i$ denotes a point drawn from a random sample with the same angular selection function as the data sample (for population $i$). We will refer to this random sample as *the random catalogue*. $D_1 D_2$, then, could be the total number of galaxy-QSO, QSO-galaxy, galaxy-star or star-galaxy pairs, to name a few constructions. If $D_1 D_2$ were a reference to the number of QSO-galaxy pairs, then $DR$ would be the number of QSO-random pairs, where the random catalogue consisted of a mock galaxy distribution. The factor $\bar{n}$ is the number of times the random catalogue is larger than the data catalogue.

Random catalogues are constructed by randomly sampling points with the same angular selection function as the sources, as discussed in Chapter 2. When estimating the correlation function in this chapter, we use a random catalogue that is 50 times bigger than the data catalogue. Note that the cross-correlation between, say, galaxies and QSOs, could be performed in one of two directions, by centring on QSOs and counting galaxies or by centring on galaxies and counting QSOs (i.e. under the exchange of labels 1 and 2 in Equation 3.5). With ideal samples, these procedures should be equivalent *provided that the angular selection of sources is accounted for by an appropriate random catalogue.* In many cases, if gradients or biases that are not accounted for in the random catalogue exist in one or both samples, the two directions may not be exactly equivalent. Henceforth, we shall drop the subscripts on $D$ and $R$ and (unless otherwise specified in the text) the reader may assume a DD or DR pair refers to a pair drawn from two different distributions.

The mechanism of calculating the cross-correlation function is to centre on one population of objects and search for the second population in bins at angle $\theta$ from the first population. The same procedure is carried out searching for objects in the random catalogue. Then $\omega(\theta)$ is derived from Equation 3.5. The scale of the measure, then, is represented by bins that are a series of annuli extending from each member of the first population. Since we are mainly interested in small-scale ($< 10$ arcminutes) cross-correlations, we

measure them locally by splitting both the NGC and SGC into 15 individual fields. The field boundaries correspond to the edges of the original APM plates from which the 2dF targets were derived (Smith et al. 1997). Correlations are then counted within each individual field, with the resulting numbers of pairs being totalled for all fields to yield the number of pairs across the entire survey area. The total number of data-data (DD) pairs and data-random (DR) pairs are then taken in ratio, as per Equation 3.5, to estimate the global correlation function. This field-to-field analysis should nullify the effects of different photometric zero-point calibrations between plates, or gross variations across strips.

More accurate estimators than Equation 3.5 have been proposed (Hamilton 1993, Landy & Szalay 1993) but purposely to deal with inaccuracies in the auto-correlation function on large scales, or more precisely on scales where the value of the correlation function is close to zero. In general, the 2QZ does not contain enough data to probe the cross-correlation function on such scales to any significance, especially given that the cross-correlation is prone to systematics present in two data sets. We will check the behaviour of our estimator on all scales when testing the accuracy of our errors.

Numerous estimates of error on the cross-correlation have been proposed. We will consider four of these. One of the simpler forms, is the Poisson error based on the number of data-data pairs in the angular bin probed:

$$\sigma_\omega^2(\theta) = \frac{1+\omega(\theta)}{DD(\theta)} \tag{3.6}$$

where we will use $\sigma_\omega$ to denote the standard error on the correlation function. Perhaps more accurate, as we will generally be measuring the significance of the correlation function compared to the null hypothesis represented by the random catalogue, is the Poisson error based on the number of data-random pairs in the angular bin probed

$$\sigma_\omega^2(\theta) = \frac{\bar{n}\,[1+\omega(\theta)]}{DR(\theta)} \tag{3.7}$$

where $\bar{n}$ is the number of times the random catalogue is larger than the data sample. We will refer to this error as the *expected Poisson error.*

The hypothesis that error on the correlation function is Poisson is not strictly fair. All else being equal, the number of counts could be highly correlated as the same points appear in different pairs that are included in many different bins, especially on large scales. Many authors (Shanks & Boyle 1994, Croom & Shanks 1996) have suggested corrections

to the Poisson form of error. Instead of corrections of this form, we will consider errors from field-to-field variations in the correlation function (see, e.g., Stevenson et al. 1985). As discussed above, our data samples are split into 30 subsamples. The cross-correlation function is measured for each subsample and the variance between the subsamples is measured. The standard error on $\omega(\theta)$ is then measured as the standard error between the susamples *inverse weighted by variance to account for different numbers of objects in each subsample*

$$\sigma_\omega^2(\theta) = \frac{1}{N-1}\sum_{L=1}^{N}\frac{DR_L(\theta)}{DR(\theta)}\left[\omega_L(\theta)-\omega(\theta)\right]^2 \tag{3.8}$$

The weighting by the number of objects in each subsample is essential, mainly because of varying completeness in the 2QZ catalogue but also as plates in the southern APM immediately East of 00h Right Ascension cover less ascension than other plates at the same ascension. We will refer to this error as *field-to-field error.*

Finally, we will measure the covariance between adjacent bins of our data. The covariance is essentially an estimate of how independent the bins are - whether bin $i$ has a tendency to take the same value as bin $j$. We might expect the covariance of the correlation function to be high, as the same data points can appear in different pairs that are counted in many different bins. The covariance as we will measure it takes the following form

$$Covar(i,j) = \sum_{M=1}^{N}\frac{\left[\omega_M(\theta_i)-\bar{\omega}(\theta_i)\right]\left[\omega_M(\theta_j)-\bar{\omega}(\theta_j)\right]}{\sigma(\theta_i)\sigma(\theta_j)} \tag{3.9}$$

Where $\theta_i$ and $\theta_j$ are two bins at different scales, and $\bar{\omega}$ and $\sigma$ represent the mean and standard deviation over a number of realisations, $M$. The covariance is 0 if the bins are independent, approaches 1 if an increase in bin $i$ leads to an increase in bin $j$ and approaches $-1$ if an increase in bin $i$ leads to a decrease in bin $j$.

To test the accuracy of both the correlation function estimator and the error associated with it, we have created 100 Monte Carlo simulations of the 2QZ QSO sample. Each simulation contains the same number of ($z > 0.4$, identification quality of "22") QSOs as the 2QZ. Each simulation has the same angular selection function as the 2QZ (as characterised by the completeness masks outlined in Chapter 2. We cross-correlate the simulated QSO samples against galaxies in groups with 7 or more members selected from the APM catalogue (for the SGC strip) and SDSS EDR (for the NGC strip). For each sample, the error estimates outlined in this section are calculated and the average over the

samples is taken. The mean value of the cross-correlation across the 100 samples is taken and the standard deviation ($1\sigma$) is calculated and recorded as the Monte Carlo error.

In the lower panel of Fig. 3.3, we display the mean cross-correlation signal across the 100 Monte Carlo simulations for the SGC QSOs (cross-correlated against galaxies in APM groups), the NGC QSOs (cross-correlated against galaxies in SDSS groups) and the combined result for the two strips. The agreement between the NGC and SGC results are excellent - better than 10 per cent of the Monte Carlo error on the NGC mean over on scales less than 20 arcminutes. The deviation of the combined result from zero, the expected result for random samples, is similarly no more than 13 per cent of the combined Monte Carlo error over all scales. We note that the shot noise would comprise 10 per cent of the Monte Carlo error (as we have 100 samples). Although the correlation function seems to begin to diverge on scales smaller than 0.8 arcminutes, the error is sufficiently large on these scales that we may consider the correlation estimator probably valid on all of the scales plotted and certainly valid on scales larger than 0.8 arcminutes. Note that the consistency of the correlation estimator across all scales indicates that the software we use to calculate the estimator is robust.

In the upper panel of Fig. 3.3 we plot the various error estimates outlined in this section. The general trend of the errors is in good agreement, although the Poisson error estimates under-predict the value of the error (as compared to the Monte Carlo estimate) on all scales. We assume that the Monte Carlo error represents a fair estimate of the true error on the correlation function. In Fig. 3.4 we plot the various errors taken in ratio to the Monte Carlo error. It is obvious that the Poisson error is an underestimate on even quite small scales and underestimates the error by as much as 40 per cent even on arcminute scales. On the smallest scales, where there are few data-data pairs on *some* of the plates, even the field-to-field estimate of error starts to diverge. This would be even more of a problem were the data samples anti-correlated (i.e. were some physical effect to further reduce the number of data-data pairs on small scales). In the following sections, we calculate a variety of cross-correlation functions using Equation 3.5 for 2QZ QSOs against galaxies in groups of APM or SDSS galaxies. We adopt the field-to-field error estimate. In Section 3.5 we shall also consider the cross-correlation of 2QZ QSOs and the groups themselves (rather than galaxies in the groups). When fitting models to these data, we will use an estimate of the error based on the variance between many Monte Carlo simulations of the data weighted to reproduce the model of interest. These weighted Monte Carlo errors are described in Section 3.5. As producing these Monte

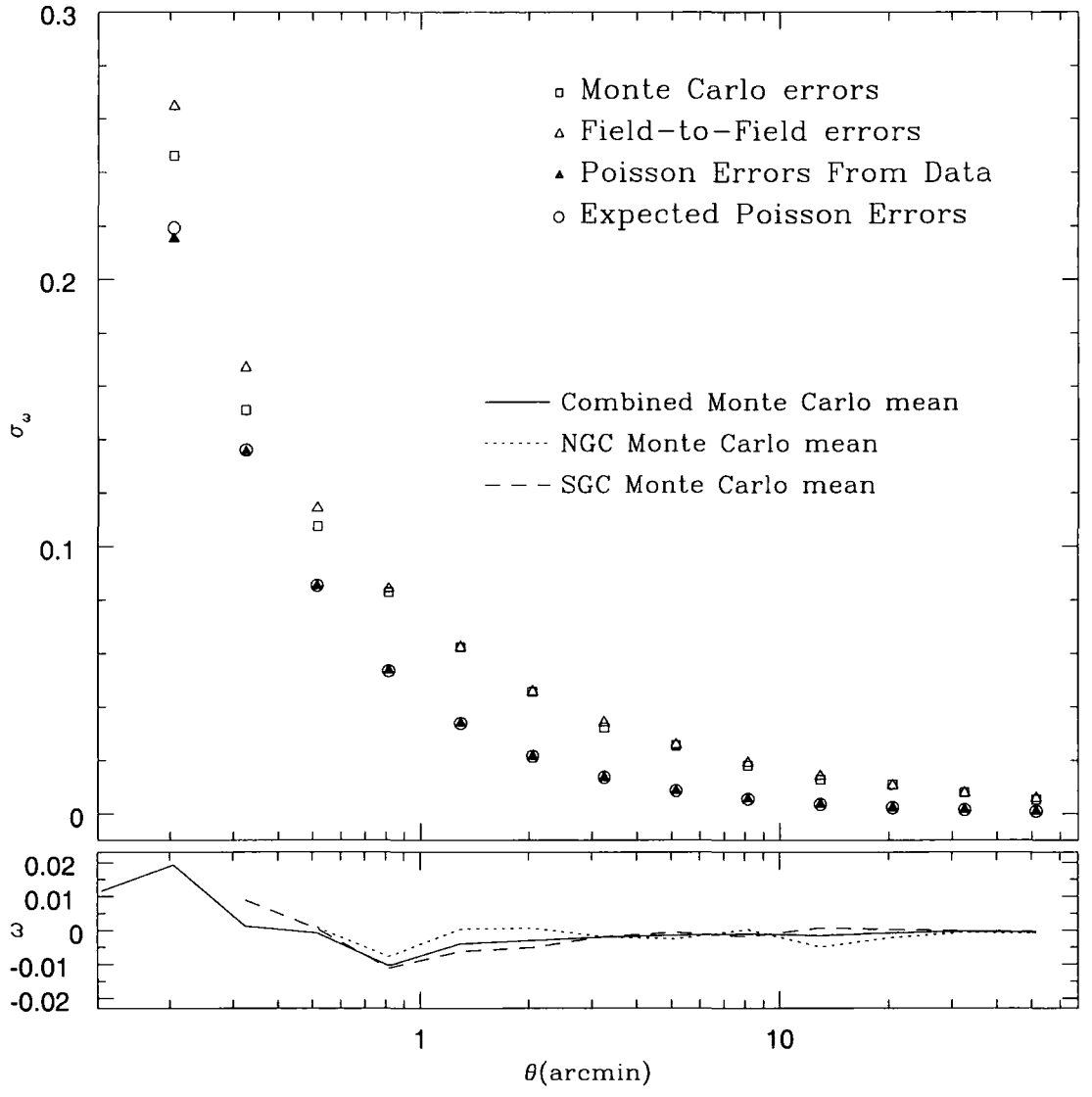


Figure 3.3: In the upper panel, we display error estimates on the correlation function for each of the equations mentioned in the text. The plotted estimates are the averages from 100 Monte Carlo simulations combined for the NGC and SGC strips. Also plotted are the Monte Carlo errors - the standard deviation of the 100 simulations. In the lower panel we plot the mean of the 100 simulations for the NGC and SGC individually and for the NGC and SGC combined.

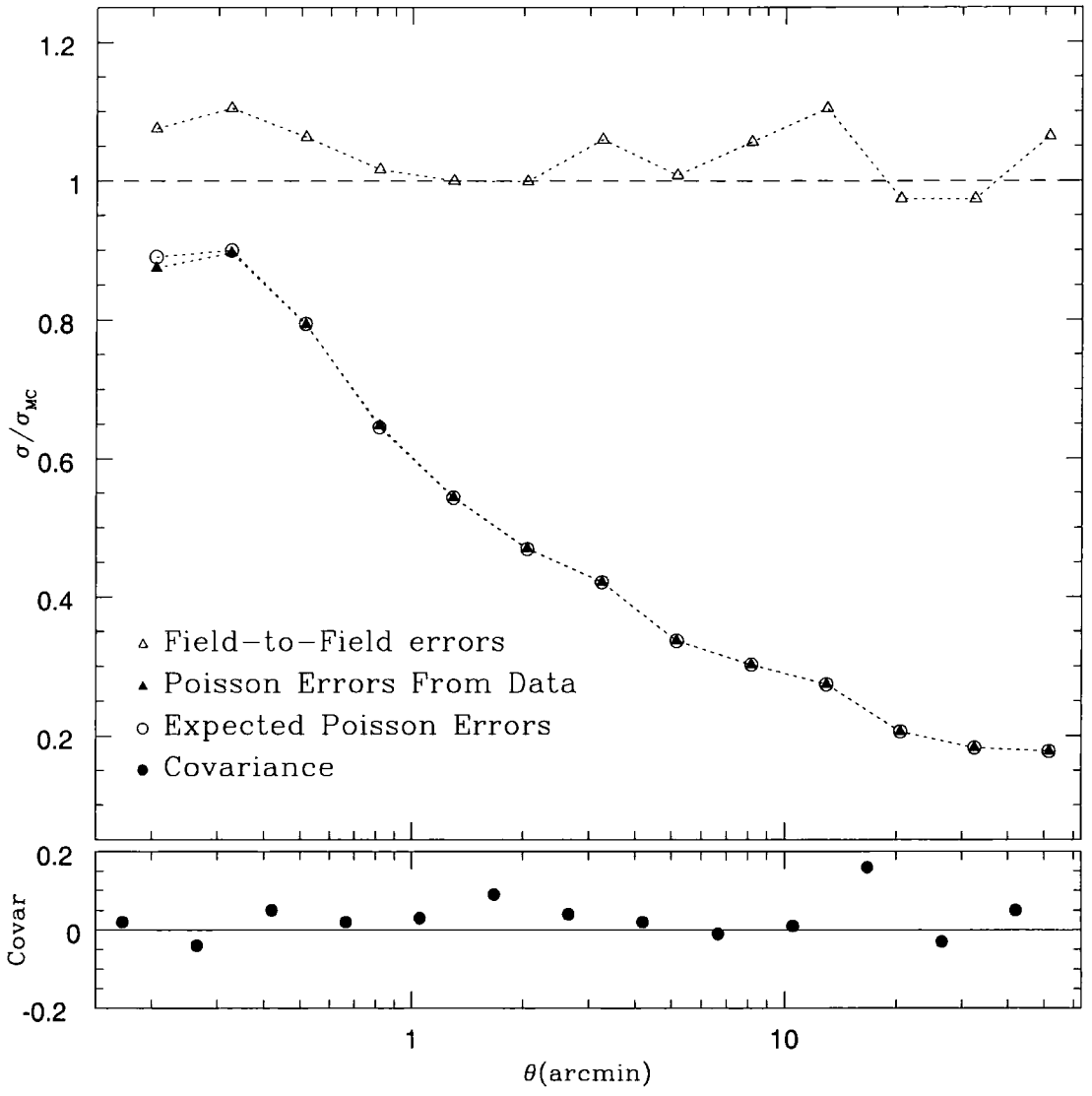


Figure 3.4: In the upper panel we display the error on the correlation function taken in ratio to the Monte Carlo error estimate, for each of the equations mentioned in the text. In all cases the errors are determined for the combined NGC and SGC sample. The dashed line at $\sigma/\sigma_{MC} = 1$ (where the Monte Carlo estimate itself would lie) is drawn for comparison. The lower panel depicts the covariance between adjacent bins determined from 100 Monte Carlo realisations

Carlo errors can be time-consuming, some figures in Section 3.5 are plotted using the field-to-field estimate of the error.

### 3.3.2 Cross-correlation of QSOs and Group Galaxies

In Fig. 3.5 (NGC and SGC) and Fig. 3.6 (combined) we display the cross-correlations between spectroscopically identifed 2QZ QSOs and galaxies in groups of at least 7 members objectively derived from the SDSS EDR in the NGC strip and the APM catalogue, in the SGC strip. In Fig. 3.6, we show both directions of correlation, we have centred on galaxies in groups and counted QSOs, and have centred on QSOs and counted galaxies in groups. Fig. 3.7 shows the comparison between our result and Boyle, Fong & Shanks (1988), where data taken from Boyle, Fong & Shanks (1988) have been scaled to allow for their projected 25 per cent contamination by stars. The results are displayed at the smallest scales for which we still believe the field-to-field errors. The largest scale displayed is about a degree - in all cases considered in this chapter the correlation function is flat on scales larger than a degree. The numbers displayed on each plot are the total data present within the 2QZ boundary. Note that the SDSS EDR sample is half the size of the 2QZ NGC strip, so does not contribute as significantly to the combined signal as the APM. A redshift cut has been made at $z = 0.4$ in the QSO sample; if the signal is caused by lensing, this will theoretically reduce the overlap between QSOs and foreground matter to at most 0.4 per cent (see Chapter 2).

Fig. 3.5 shows that there is good consistency between the NGC and SGC correlation functions for QSOs versus galaxies in groups, justifying combining the signals. Fig. 3.6 demonstrates that there is excellent consistency in the cross-correlation signal between QSOs and group galaxies irrespective of the direction in which the function is calculated, suggesting the signal is robust, free from the influence of any gradient or incompleteness in the samples used (and that the 2QZ masks discussed in Chapter 2 have been adequately constructed). There is a significant $2.9\sigma$ anti-correlation between galaxy groups and spectroscopically identified 2QZ QSOs on scales out to 10 arcminutes, based on collecting the data in a single 10 arcminute bin and calculating the *rms* field-to-field variation. For an Einstein-de-Sitter cosmology, 10 arcminutes would translate to 1 $h^{-1}$ Mpc at an average group redshift of $z \sim 0.15$ (1.1 $h^{-1}$ Mpc for $\Lambda$CDM . The average anti-correlation in such a 10 arcminute bin is $-0.049$. Our data compare well in Fig. 3.7 with the anti-correlation discovered by Boyle, Fong & Shanks (1988) when correlating a UVX object sample with cluster galaxies. Boyle, Fong & Shanks (1988) declared a more significant

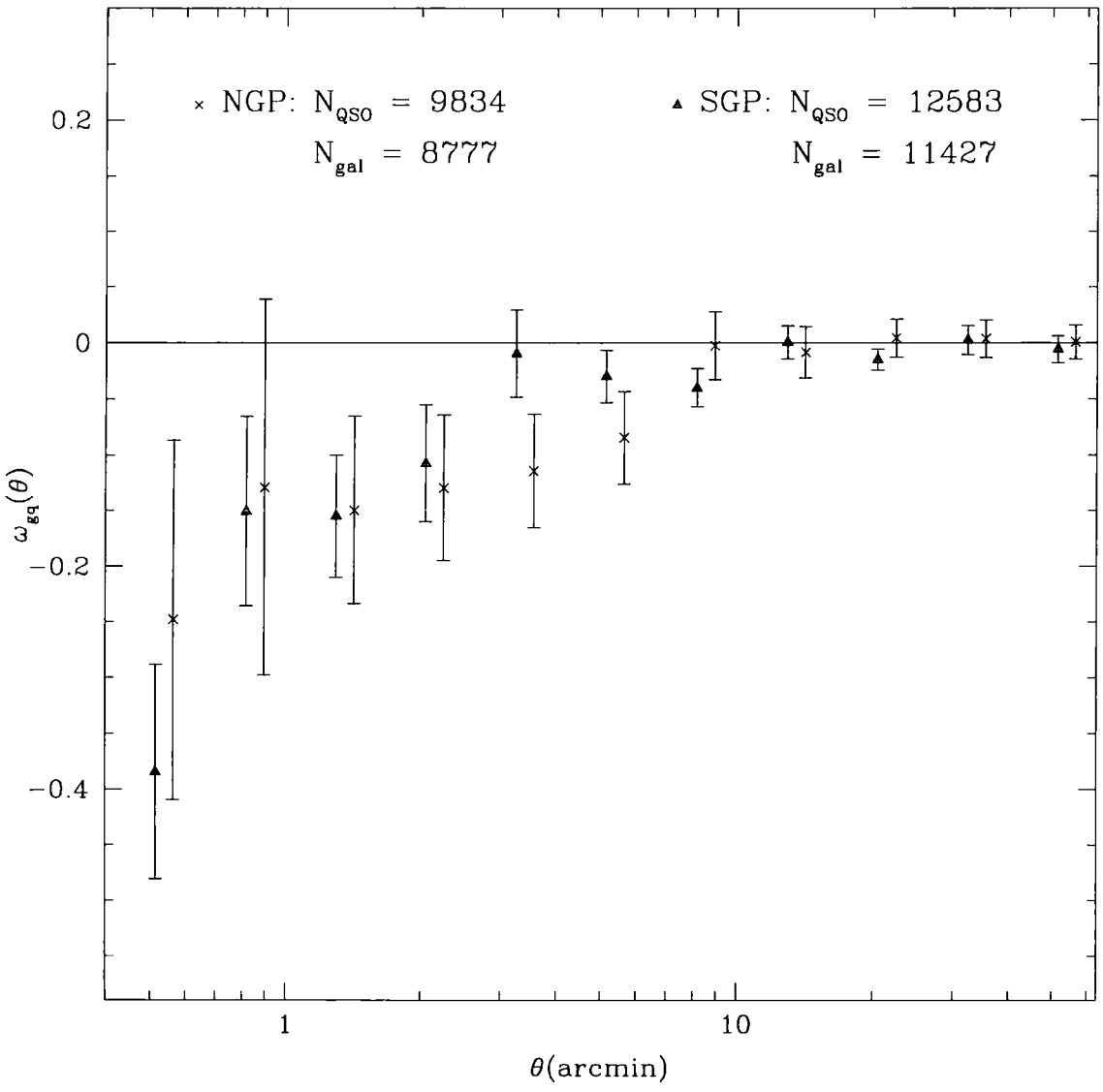


Figure 3.5: The cross-correlation between 2QZ QSOs, and galaxies of limiting magnitude $b_J = 20.5$ found in objectively derived groups of at least seven members, for both 2QZ strips. Crosses correspond to the NGC strip, triangles to the SGC strip. The numbers of each sample within the boundaries of the 2QZ are displayed. Errors are field-to-field.

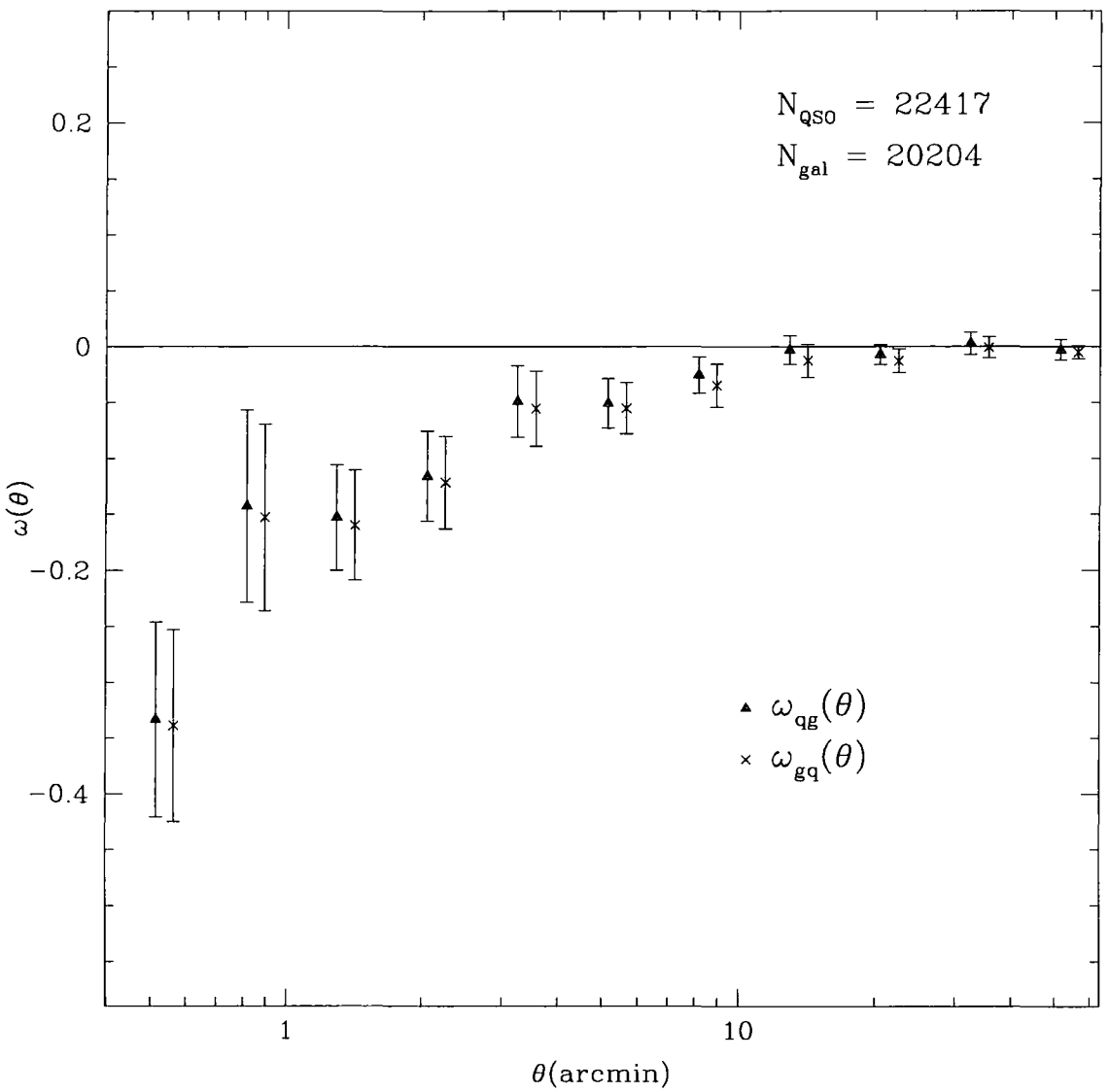


Figure 3.6: The cross-correlation between galaxies of limiting magnitude $b_J = 20.5$ found in objectively derived groups of at least seven members and 2QZ QSOs, combined for both the southern and northern 2QZ strips. Both results centring on QSOs and counting galaxies (triangles) and centring on galaxies and counting QSOs (crosses) are presented. The numbers of each sample within the boundaries of the 2QZ are displayed. Errors are field-to-field.

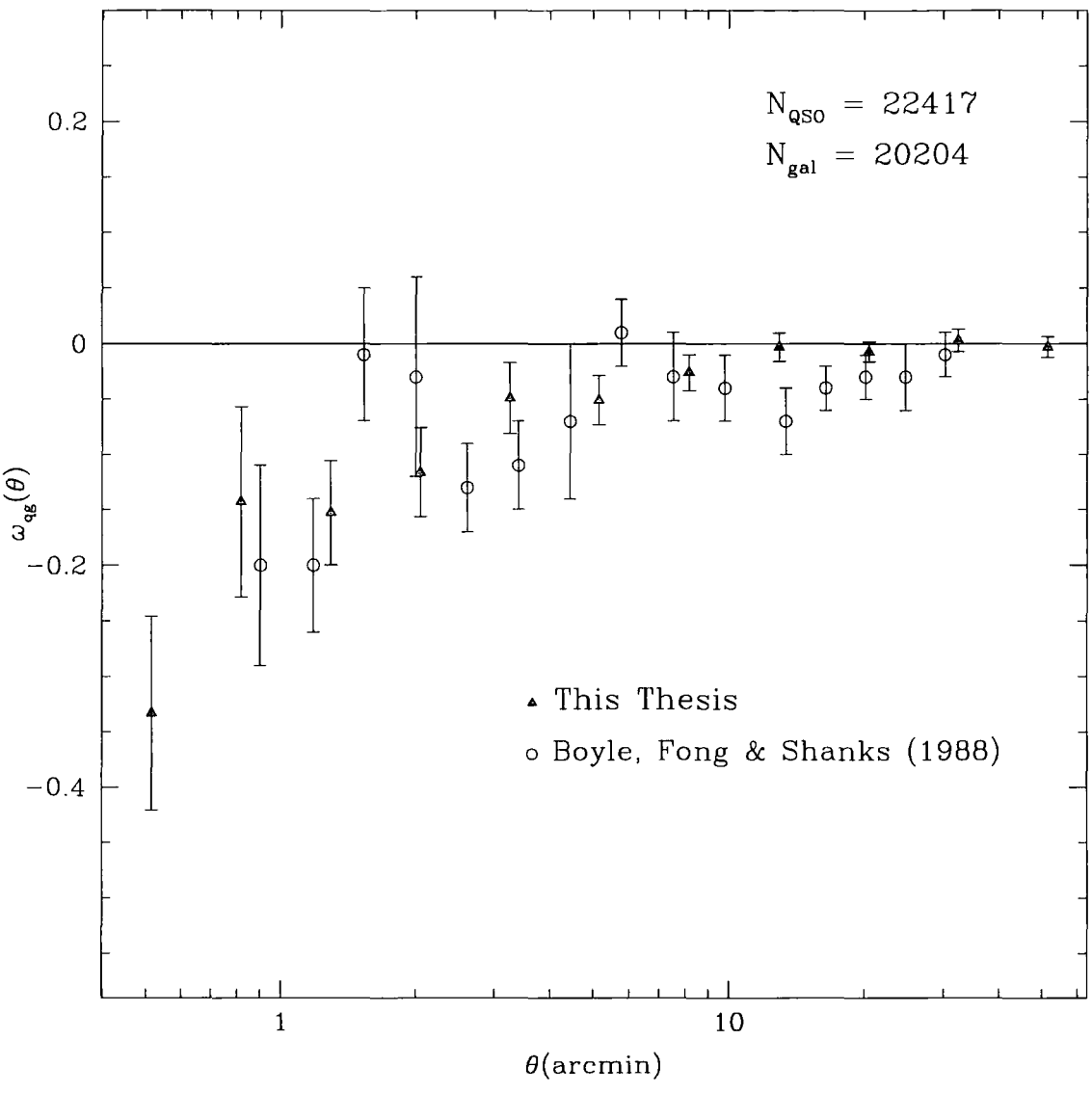


Figure 3.7: The comparison between Fig. 3.6 and the result of Boyle, Fong & Shanks (1988). The data from Boyle, Fong & Shanks (1988) (open circles) have been scaled to account for their stellar contamination.

$4\sigma$ signal on <4 arcminute scales.

To test the anti-correlation between group galaxies and QSOs is not a systematic of the 2dF observing process, we need a control sample that has been through the same measurement and reduction as the QSO sample. Boyle, Fong & Shanks (1988) found no correlation between a large ($\sim$ 27,000) control sample of non-UVX stars and group galaxies, however, our situation is notably more complicated. As we require a control sample that has been through the same observing processes as the QSOs, including spectroscopic identification, we are restricted to a smaller sample of stars ($\sim$ 12,000) compared to our QSO sample ($\sim$ 22,500). Additionally, the selection criterion of the 2QZ tends to pick out specific, UV-bright populations of stars, such as White Dwarfs.

In Fig. 3.8 (NGC and SGC) and Fig. 3.9 (combined) we show the cross-correlation between spectroscopically identified 2QZ stars and galaxies in groups. Unlike the QSO result, the NGC and SGC cross-correlations are not entirely consistent, suggesting that the different physical distribution of stars in the two strips effects the correlation function. The combined results are also inconsistent, although the errors are sizable. The combined results show evidence of gradients on large scales, as, unlike in the case of the QSOs, the cross-correlation result is dependent on whether we centre on stars and count galaxies or centre on galaxies and count stars. Indeed, the distribution of stars within our galaxy has a gradient with galactic latitude in both the NGC and SGC, unlike the distribution of QSOs, which, being cosmological, has no large-scale gradient. The cross-correlation between stars and galaxies has some negative points on 3-7 arcminute scales but does not have the same form as the QSO anti-correlation. In particular, the star-signal does not continue to decrease on scales less than 3 arcminutes. If we calculate the significance of the signal out to 10 arcminutes, we find a $1.4\sigma$ anti-correlation for the result centring on stars and counting galaxies and a $2.9\sigma$ anti-correlation for the result centring on galaxies and counting stars. Much of this discrepancy is caused by the large-scale gradients; if the galaxy-centred result is moved up so there is no anti-correlation on large scales, it comes into line with the star-centred result. The anti-correlation in both of the combined star results is caused entirely by the few points on 3-7 arcminute scales. Surprisingly, most of the anti-correlation between stars and group galaxies is actually caused by White Dwarfs. If we discard the White Dwarfs from the stellar sample, there is no significant anti-correlation between stars and group galaxies (see Fig. 3.10). We discuss a possible physical explanation of this effect in Chapter 5 but the most likely explanation for the anti-correlation seen at 3-7 arcminutes in both the stars and White Dwarfs is a statistical

fluctuation due to the low numbers of these objects.

We regard the results using the above control samples of stars as encouraging in terms of ruling out a systematic source for the QSO-group anti-correlation. However, because of the low numbers of stars, their gradients in Right Ascension and the anti-correlation detected at 3-7 arcminutes, stars may not form the ideal control sample and there may be a residual doubt as to whether there is a systematic contribution to the QSO anti-correlation on 3-7 arcminute scales caused by the fibre positioning constraint of the 2dF instrument. After all, the 2dF Survey observed $b_J < 19.5$ galaxies and QSOs simultaneously and in dense fields close pairs of objects may have been missed due to the minimum fibre spacing, even though 2dF candidates were given a higher priority in the fibre allocation to prevent imprinting the galaxy structure on the QSO distribution. It is easy to show there is no fibre positioning effect by comparing the cross-correlation of the 46,000 objects targeted as 2QZ QSO candidates in the 2dF input catalogue (Smith et al. 1997) with the 43,000 objects that were observed spectroscopically. These results are shown respectively in the upper left-hand and lower left-hand panels of Fig. 3.10. These two results are in all respects identical with no systematic difference between them, leading us to conclude that there is no anti-correlation induced on these scales by the fibre positioning constraint. Therefore we conclude that the QSO-group galaxy anti-correlation is probably real.

The reason, we note, that the correlation between group galaxies and the input catalogue is flat, even though it hides the significantly anti-correlated QSO signal, becomes apparent when we split the input catalogue up into its constituent parts. The right-hand four panels in Fig. 3.10 display the main subsamples of the input catalogue. We can see that the Narrow Emission Line Galaxies included in the input catalogue exhibit a positive correlation that cancels out the anti-correlation exhibited by QSOs.

### 3.3.3 The dust hypothesis

Boyle, Fong & Shanks (1988) originally attributed the anti-correlation between QSOs and galaxies in groups to interim dust in clusters, finding that an absorption in the $B-$band of $A_B \sim 0.2$ mag was sufficient to cause their observed anti-correlation. Such absorption is at the upper limit allowed by Ferguson (1993), who quoted a maximum reddening of $E(B-V) \leq 0.06$ from a composite study of the $Mg_2$ index of 19 nearby clusters and rich groups. Using colour information provided by the 2dF survey, we can limit the culpability of dust in causing the anti-correlation signal.

In Fig. 3.11 we show the distribution of colours of 2QZ QSOs that lie within 10 arcmin-

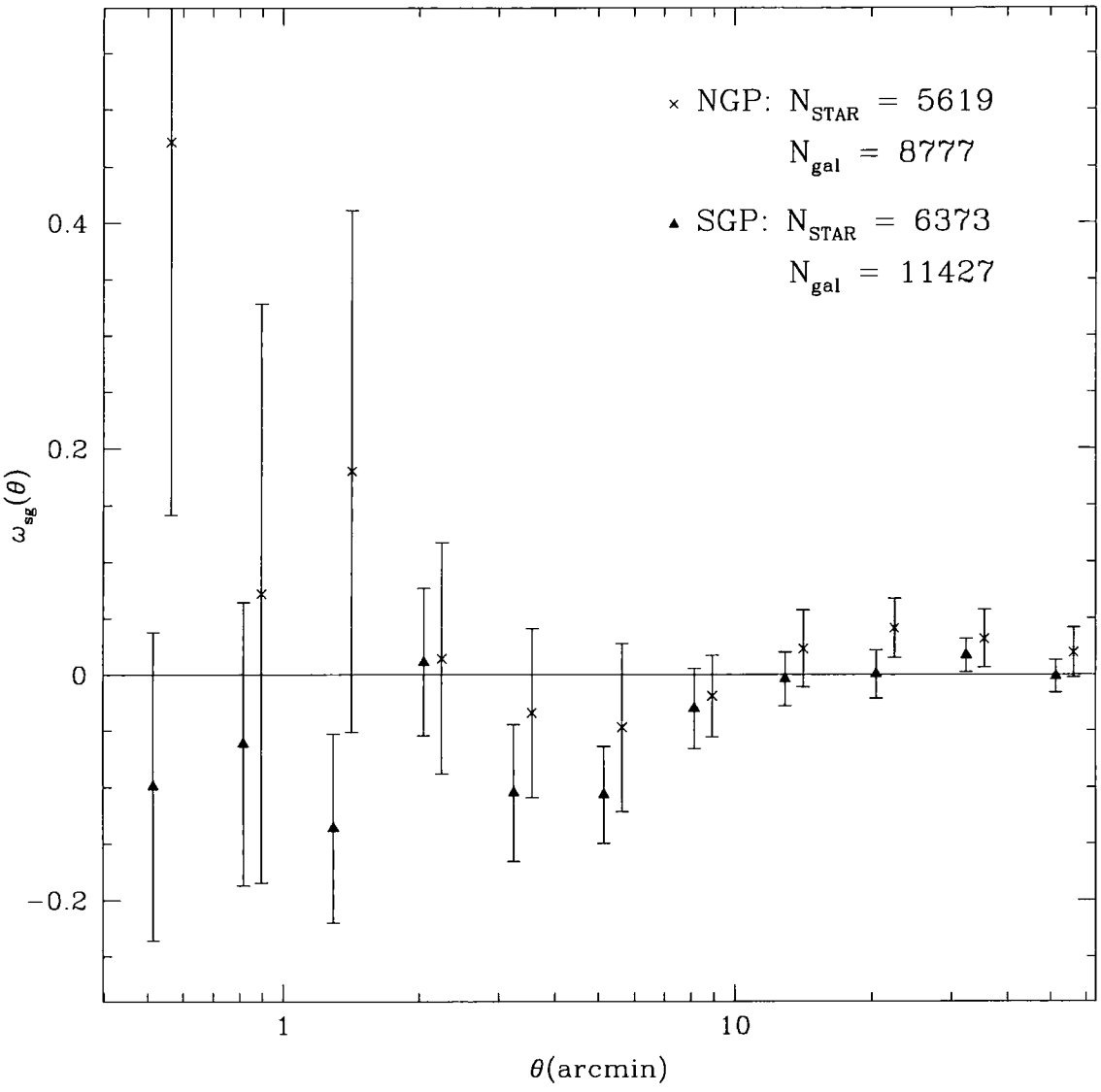


Figure 3.8: The cross-correlation between 2QZ stars, and galaxies of limiting magnitude $b_J$ = 20.5 found in objectively derived groups of at least seven members, for both 2QZ strips. Crosses correspond to the NGC strip, triangles to the SGC strip. The numbers of each sample within the boundaries of the 2QZ are displayed. Errors are field-to-field.

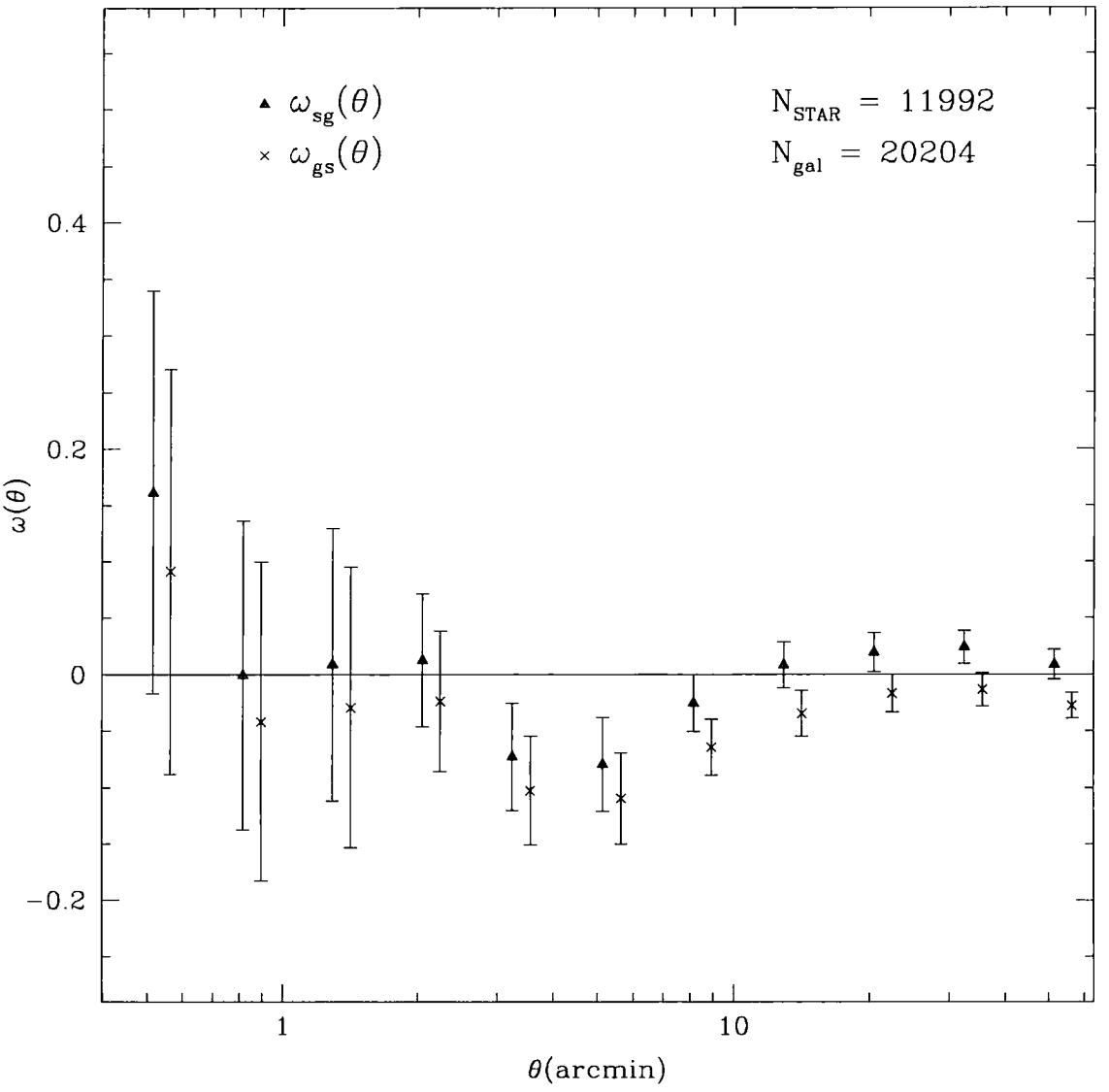


Figure 3.9: The cross-correlation between galaxies of limiting magnitude $b_J = 20.5$ found in objectively derived groups of at least seven members and 2QZ stars, combined for both the SGC and NGC 2QZ strips. Both results centring on stars and counting galaxies (triangles) and centring on galaxies and counting stars (crosses) are presented. The numbers of each sample within the boundaries of the 2QZ are displayed. Errors are field-to-field.

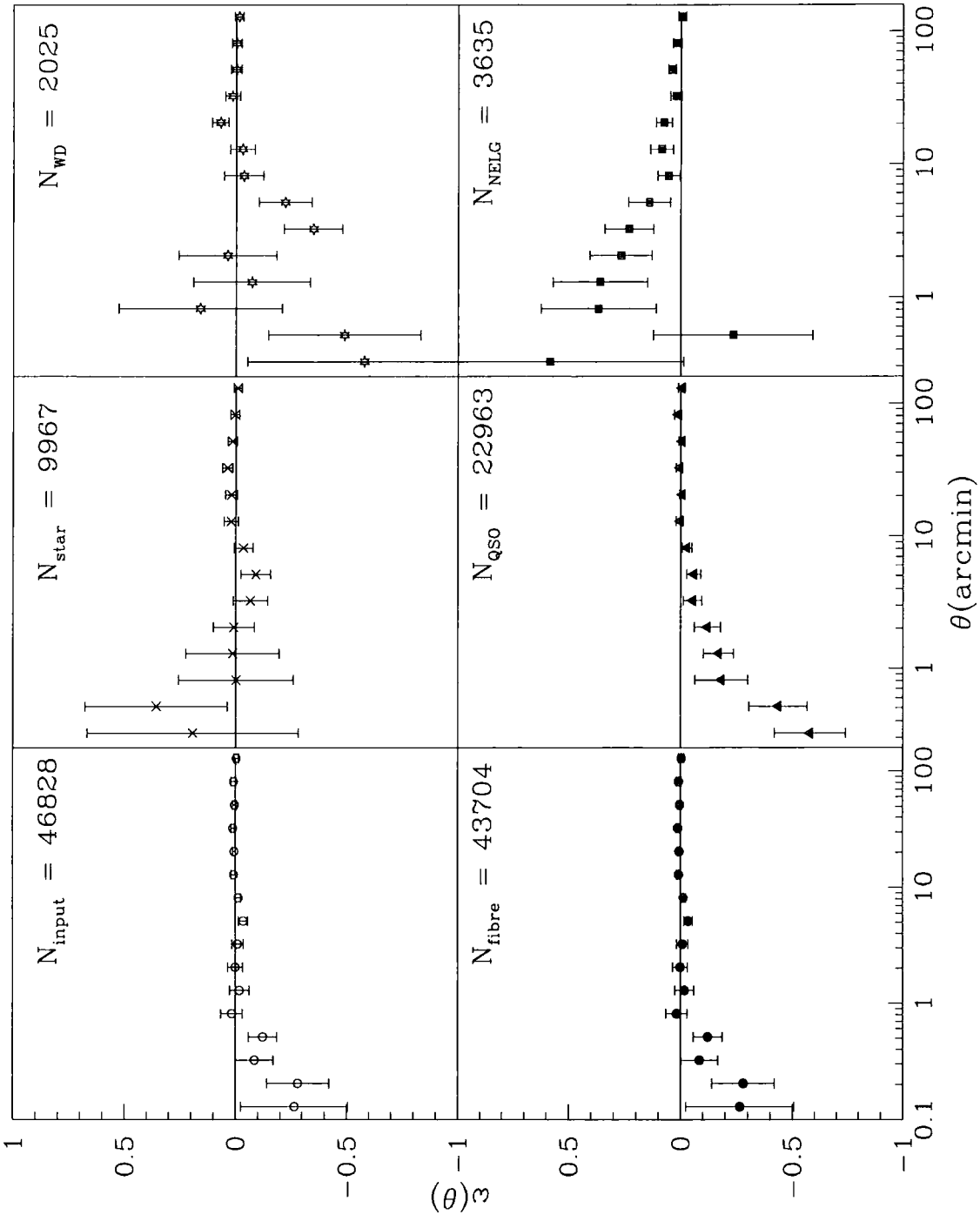


Figure 3.10: The cross-correlation between 20204 galaxies to $b_J < 20.5$ found in groups of seven or more members and 2QZ objects, combined for NGC and SGC 2QZ strips. The upper-left panel is the 2QZ input catalogue, every QSO candidate targeted in the preparation of the 2QZ photometric sample. The lower-left panel is the fibre catalogue, every QSO candidate observed with a fibre. Other panels are for objects that comprise most of the fibre catalogue; stars, QSOs, White Dwarfs and NELGs. The numbers of each subsample are displayed. Errors are field-to-field.

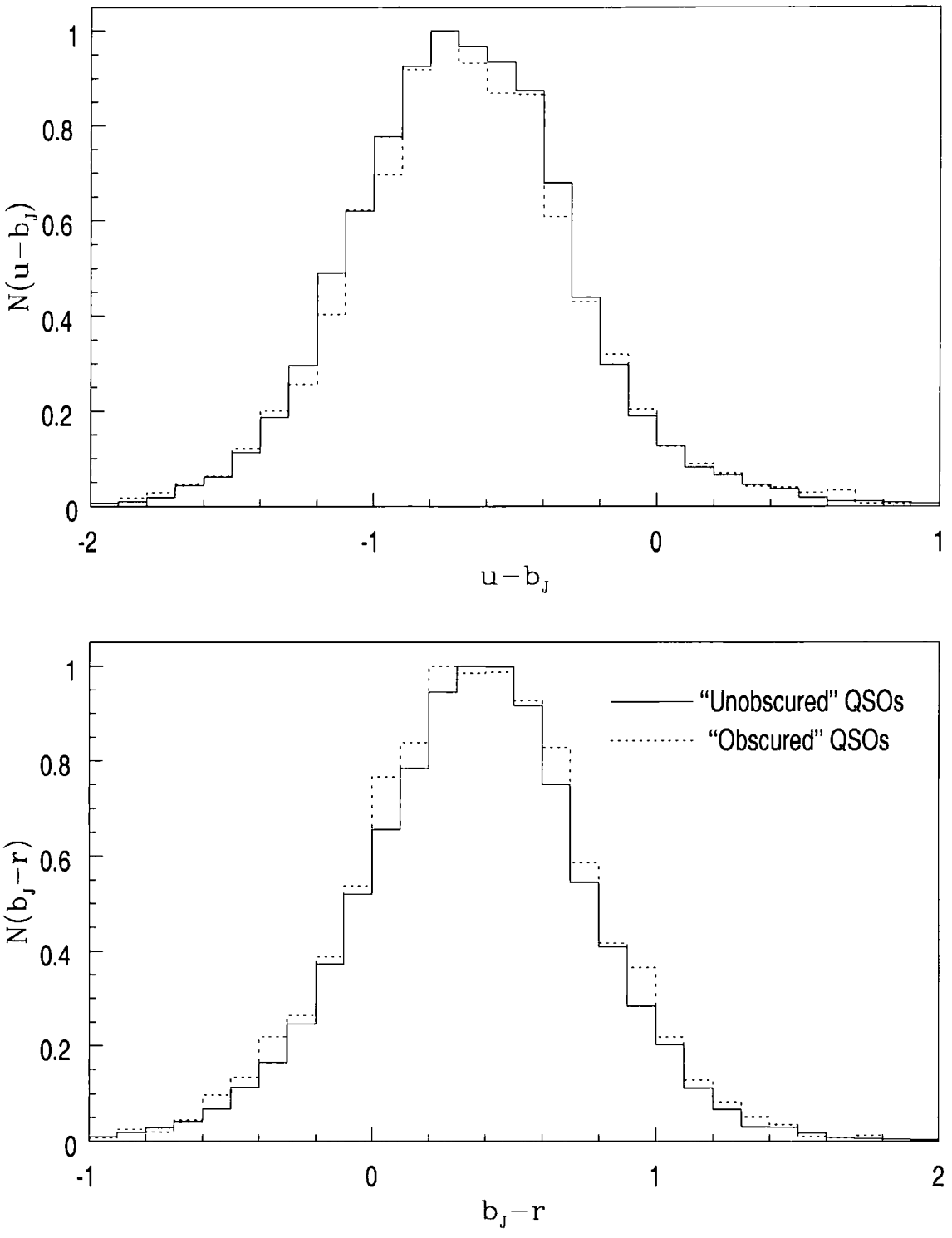


Figure 3.11: The relative distribution of colours of 2QZ QSOs by proximity to a galaxy group. The solid histogram shows the colours of QSOs that *do not* lie within 10 arcminutes of a group centre. The dashed histogram shows the colours of QSOs that lie within 10 arcminutes of a group centre.

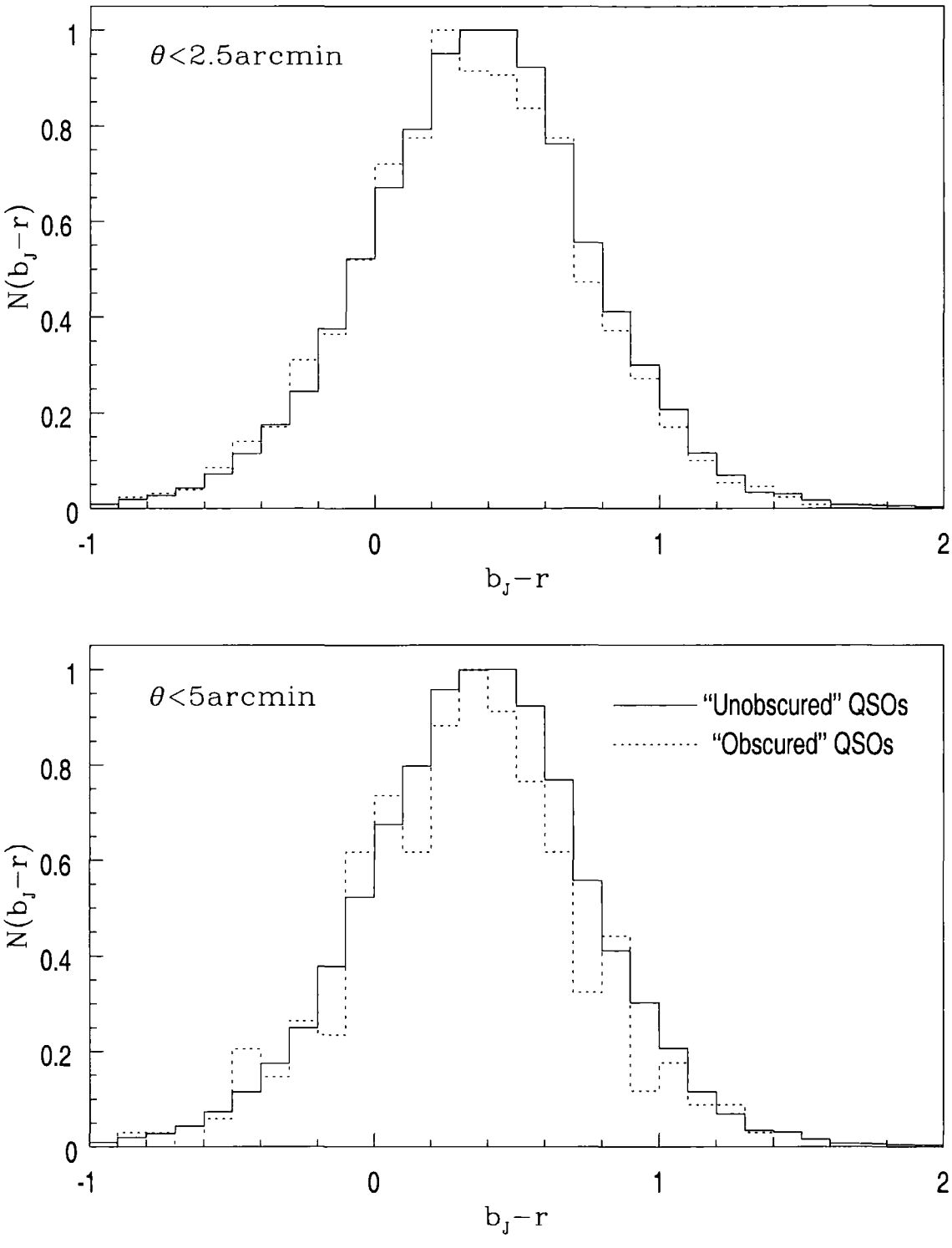


Figure 3.12: The relative distribution of colours of 2QZ QSOs at different angular radii from galaxy groups. In the upper panel, the solid histogram shows the colours of QSOs that *do not* lie within 2.5 arcminutes of a group centre (5 arcminutes for the lower panel) and the dashed histogram shows the colours of QSOs that lie within 2.5 arcminutes of a group centre (5 arcminutes for the lower panel).

utes of any group centre ("obscured" QSOs) compared to that of QSOs that do not lie within 10 arcminutes of any group centre ("unobscured" QSOs). We have selected 10 arcminutes as the radius of interest as it corresponds to the extent of the anti-correlation signal in Fig. 3.5. Were the anti-correlation due to intervening dust in galaxy groups, as proposed by Boyle, Fong & Shanks (1988), we might expect to see a complementary reddening of QSOs on the scale of the anti-correlation, and the distributions of "obscured" and "unobscured" QSOs would differ.

A two-sample Kolmogorov-Smirnov test fails to reject the null hypothesis that the "obscured" and "unobscured" distributions are drawn from the same parent population, for both colours. Student's t-test demonstrates that the means of the "obscured" and "unobscured" samples are in agreement. For the $u - b_J$ colour distributions, the "unobscured" mean and standard error are $-0.6789 \pm 0.0030$ and for the "obscured" are $-0.6687 \pm 0.0064$. For the $b_J - r$ colour distributions, the "unobscured" mean and standard error are $0.3644 \pm 0.0030$ and the "obscured" mean and standard error are $0.3626 \pm 0.0065$. There are 4025 "obscured" QSOs and 17752 "unobscured" QSOs. Student's t-distribution sets the following 95 (99) per cent upper limits on reddening between the two distributions: $E(u - b_J) = 0.012$ ($E(u - b_J) = 0.016$), $E(b_J - r) = 0.012$ ($E(b_J - r) = 0.016$).

Our average group size is around 2.5 arcminutes and larger groups in our sample have an angular size of 5 arcminutes, so we might also be interested in any reddening on these scales. Fig. 3.12 displays the distribution of $b_J - r$ colours for "obscured" and "unobscured" QSOs on these scales. Repeating the statistical analyses on these scales, Student's t-test suggests the 95 (99) per cent reddening limits between the QSO population within 5 arcminutes of any group centre and the QSO population lying greater than 10 arcminutes are $E(b_J - r) = 0.020$ (0.028) and the 95 (99) per cent reddening limits between the QSO population within 2.5 arcminutes of any group centre and the remainder are $E(b_J - r) = 0.039$ (0.056). These reddening limits apply for both the $u - b_J$ and $b_J - r$ colours. The limits inevitably increase as the "obscured" population dwindles in size, although there are still 286 QSOs within 2.5 arcminutes of any group centre. Again, for both the 2.5 arcminute and 5 arcminute scales, the Kolmogorov-Smirnov test failed to reject the hypothesis that the "obscured" and "unobscured" distributions are drawn from the same parent distribution of colours. Assuming the Galactic dust law adopted by Schlegel, Finkbeiner & Davis (1998), the 95 per cent limit from the $b_J - r$ colours on B-band absorption within 5 arcminutes ($\sim$ 0.5 $h^{-1}$ Mpc) of any group centre is $A_B < 0.06$ mag. The similar limit within 2.5 arcminutes of any group centre is $A_B < 0.13$ mag.

A simple model (see Boyle, Fong & Shanks 1988, or Section 5.4.4.1 for details), taking the slope of the QSO number-magnitude counts to be 0.29, suggests that these levels of absorption correspond to a correlation function of $\omega_{cq} < -0.039$ for $A_B < 0.06$ mag and $\omega_{cq} < -0.077$ for $A_B < 0.13$ mag. For the entire sample out to 10 arcminutes, the 95 per cent reddening limits suggest an anti-correlation of only $-0.029$. The actual values of the detected anti-correlation on these scales are $\omega_{cq}(\theta < 2.5') = -0.15 \pm 0.04, \omega_{cq}(\theta < 5') = -0.09 \pm 0.03, \omega_{cq}(\theta < 10') = -0.049 \pm 0.017$.

It seems that 2QZ QSOs close to galaxy groups are insufficiently reddened to explain the anti-correlation signal as an effect of dust in galaxy groups. It might be argued that as the 2QZ is UVX-selected, our reddening values are biased by the limits set on the colours of 2QZ QSOs. We do not mean to argue that our values are objective determinations of the extent of dust in galaxy groups, only that there is insufficient reddening of QSOs close to galaxy groups within the 2QZ to explain the anti-correlation signal. Undoubtedly, there will be some heavily reddened QSOs close to group centres that the 2QZ fails to observe, however, the QSOs the survey *does* observe have no tendency to redness close to group galaxies. The 2QZ sample colour-colour distribution peaks significantly bluewards of any colour limit, so we believe that the low reddening measure is not forced by the UVX-selection. We have transformed colours to absorption limits assuming the dust law from Schlegel, Finkbeiner & Davis (1998) for the Milky Way galaxy. Other galaxies have been shown to have slightly different dust laws (Calzetti, Kinney & Storchi-Bergmann 1994, Bolzonella, Miralles & Pelló 2000), notably, some galaxies have been shown to have "greyer" dust laws (meaning more absorption by dust for a given change in colour). In Section 5.4.4.1 we will discuss this at some length, and show that if galaxies or groups obscuring QSOs have a Calzetti, Kinney & Storchi-Bergmann (1994) dust law, we would unserestimate the amount our QSO samples were being obscured by 25 per cent, which is still insufficient to explain the anti-correlation between 2QZ QSOs and galaxies in groups.

Previous evidence for dust in galaxy groups has been controversial, with some authors claiming detections and others claiming upper limits. Girardi et al. (1992) confirm galaxies in local groups that are blueshifted relative to the group average tend to have a larger colour excess than that group average, and suggest background galaxies falling towards the group centre suffer reddening by dust in the group. The amount of dust Girardi et al. (1992) infer is $E(B-V) \sim 0.1 - 0.2$ on 0.75 Mpc scales, equivalent to an absorption of $A_B \sim 0.4 - 0.8$ mag (for the usual galactic dust law). Such an amount of absorption is at

odds with the upper limit quoted by Ferguson (1993). Ferguson studied the correlation of $Mg_2$ index with $B-V$ colour for local groups and clusters, and compared these results to a sample of field galaxies. Ferguson (1993) quoted $E(B-V) \leq 0.06$ as a conservative (90 per cent) upper limit on reddening in a sample of clusters and groups and, similarly, an upper limit of $E(B-V) \leq 0.05$ ($A_B < 0.2$ mag) for a sample of poorer groups. Ferguson (1993) considers a scale of 0.5 Mpc to be the central contribution to his reddening limits. The 95 per cent upper limits on our reddening results out to 0.5 $h^{-1}$ Mpc (5 arcminutes) allow for us much absorption by dust as $A_B < 0.06$ mag, somewhat more consistent with the results of Ferguson (1993) rather than those of Girardi et al. (1992). Our results are particularly consistent with a recent paper by Nollenberg, Williams & Maddox (2003), who find 99 per cent limits on reddening by dust in APM groups of $E(B-V) \lesssim 0.02$ (using a similar technique to that used in this section, to limit the extent that background *galaxies* could be reddened by foreground groups).

The possibility remains that some specifically tailored dust model could explain the anti-correlation. One possibility (which we discuss further in Chapter 5) is a smooth distribution of dust around galaxy groups that does not obey the reddening laws observed in our galaxy (i.e. particularly "grey" dust; Calzetti, Kinney & Storchi-Bergmann 1994). A slightly less ad hoc explanation is a patchy distribution of dust around galaxy groups that is heavily concentrated in the line of sight direction, such that QSOs are completely obscured without being reddened. Averaged over many groups, this would have the effect of removing QSOs from the sample near galaxy groups without overtly reddening the QSO sample on similar scales. Finally, we must consider the possibility that multiple effects of dust and lensing combine to produce the observed anti-correlation. So, although our current analysis can rule out a smooth distribution of typical dust around galaxy groups, we cannot deny more tailored dust models, such as grey dust, patchy dust or a combination of gravitational lensing and dust.

## 3.4 Modelling the Statistical lensing of QSOs

We have seen that the $3\sigma$ anti-correlation between 2QZ QSOs and galaxies in groups is unlikely to be either a selection effect or a product of normal, smoothly distributed dust in galaxy groups obscuring QSOs. In this section we consider the possibility that the anti-correlation results from the statistical lensing of QSOs by foreground galaxy groups.

### 3.4.1 Lensing

As Einstein (Einstein 1915) noted (see also Appendix C), in a spherically-symmetric geometry, a mass $M$ will bend a ray of light passing at impact parameter $\xi$ through an angle $\alpha_{\rm bend}$ (Schneider, Ehlers & Falco 1992)

$$\alpha_{\rm bend} = \frac{4GM(<\xi)}{\xi c^2} = \frac{D_{\rm s}}{D_{\rm ls}}(\theta - \theta_{\rm q}) \tag{3.10}$$

where $D_{\rm l}$, $D_{\rm s}$ and $D_{\rm ls}$ are the angular diameter distances of the lens as measured by the observer, the source as measured by the observer and the source as measured by the lens, respectively; $\theta$ is the angle between the image, the observer and the centre of the lens; $\theta_{\rm q}$ is the angle between the source, the observer and the centre of the lens. Surface brightness conservation implies each image will be magnified by a factor

$$\mu = \left| \frac{\theta}{\theta_{\rm q}} \frac{{\rm d}\theta}{{\rm d}\theta_{\rm q}} \right| \tag{3.11}$$

(Turner, Ostriker & Gott 1984). Lensing influences a sample of background objects in two competing ways. Fainter objects are lensed into a magnitude limitied sample, increasing the number density of that sample but the area behind the lens is proportionately expanded, reducing the sample's number density. Narayan (1989) quantified this effect as a 'net enhancement factor'

$$q = \frac{1}{\mu} \frac{N(< m + 2.5 \log \mu)}{N(< m)} \tag{3.12}$$

The Number-magnitude relation can be approximated as a power law, with $N(< m) \propto 10^{\alpha m}$ (Boyle, Fong & Shanks 1988), allowing us to express the enhancement factor as

$$q = \frac{1}{\mu} \frac{10^{\alpha(m+2.5\log\mu)}}{10^{\alpha m}} = \mu^{2.5\alpha - 1} \tag{3.13}$$

(Croom 1997). Note that there is no real need to make this approximation - but we shall show later that interpolating across the full form of $N(< m)$ and making this approximation yield similar results at the limit of the 2QZ. Now, the net enhancement factor is the ratio of the observed (lensed) flux and the true (unlensed) flux. The correlation function $\omega(\theta)$ may be expressed as the ratio of observed pairs of objects to expected pairs

of objects. Typically, 1 is subtracted from the correlation function to account for the expected normal background of pairs. Hence:

$$\omega(\theta) = q - 1 = \mu^{2.5\alpha - 1} - 1 \tag{3.14}$$

Equation 3.14 dictates $\omega(\theta) = 0$ when $\alpha = 0.4$. For higher values of $\alpha$ we would observe a correlation, and for lower values, an anti-correlation. Thus, the lensing effect is dependent on the slope of the number-magnitude relation. Constraints on the slope of $N(< b_J)$ were derived in Section 2.3.

### 3.4.2 Dark matter profiles

The Singular Isothermal Sphere has a mass surface density defined in terms of its radial velocity dispersion $\sigma(r)$ (Schneider, Ehlers & Falco 1992)

$$\Sigma_{SIS} = \frac{\sigma^2}{2Gr} \tag{3.15}$$

which can be integrated out to $r = \xi$, the impact parameter as defined for Equation 3.10, to find

$$M = \int_0^{\xi} 2\pi r \Sigma_{SIS} dr = 2\pi \int_0^{\xi} \frac{\sigma^2}{2Gr} r dr = \frac{\pi \sigma^2 \xi}{G} \tag{3.16}$$

which can be substituted into Equation 3.10, and then Equation 3.11 to determine that the SIS will magnify background sources by the factor

$$\mu = \left| \frac{\theta}{\theta - 4\pi \frac{D_{ls}}{D_s} (\frac{\sigma}{c})^2} \right| \tag{3.17}$$

with $c$ being the speed of light and $\sigma$ the velocity dispersion of the SIS. The term $4\pi \frac{D_{ls}}{D_s} (\frac{\sigma}{c})^2$ in the denominator is the "Einstein Radius". The Einstein Radius essentially marks the boundary between the weak and strong lensing regimes (see Appendix C). Equation 3.14 and Equation 3.17 can be combined to constrain the predicted form of the correlation function from magnification through an SIS profile.

N-body simulations have provided a universal (NFW) density profile for dark matter haloes (Navarro, Frenk & White 1995, Navarro, Frenk & White 1996, Navarro, Frenk & White 1997) that has been independently observationally confirmed with some success

(Bartelmann et al. 1998, Thomas et al. 1998). We have also constructed a simple model to determine the form of the correlation function based on lensing through such a halo.

The NFW profile in the form

$$\rho(r) = \frac{\delta_c \rho_c}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2} \tag{3.18}$$

where $r_s$ is a representative radial scale and $\rho_c$ is the critical density or the universe at the redshift of the dark matter, appears to be a reasonable description of haloes in ΛCDM simulations spanning 9 orders of magnitude in mass, from globular clusters to galaxy clusters. Ideally, we would wish to study statistical lensing utilising just such a realistic density profile. Following Bartelmann (1996), we rewrite the profile as

$$\rho(x) = \frac{\rho_s}{x\,(1+x)^2} \tag{3.19}$$

and consider lensing around the profile at a radius of the impact parameter $\xi$, meaning that $x$ is defined by

$$x = \frac{\xi}{r_s} = \frac{D_l \theta}{r_s} \tag{3.20}$$

where $D_l$ and $\theta$ are defined as for the SIS, above. One can then combine Equations 3.11 and 3.20 to yield

$$\mu = \left| \left(\frac{r_s}{D_l}\right)^2 \frac{x}{\theta_q} \frac{\mathrm{d}x}{\mathrm{d}\theta_q} \right| \tag{3.21}$$

Now, Maoz et al. (1997) determine the characteristic scale $r_s$ from an empirical fit to Figure 9 of Navarro, Frenk & White (1997) as

$$r_s = 300 \left(\frac{M}{10^{15} \mathrm{M}_\odot}\right)^{\gamma} h^{-1} \mathrm{kpc} \tag{3.22}$$

where $h$ is the Hubble constant in units of 100 km $\mathrm{s}^{-1}\mathrm{Mpc}^{-1}$ and $\gamma$ varies between cosmologies (for CDM, $\gamma \sim \frac{1}{3}$). Maoz et al. (1997) go on to rewrite Equation 3.10 as

$$\alpha_{\mathrm{bend}} = \frac{4GM_{1.5}}{c^2 r_s g(1.5Mpc/r_s)} \frac{g(x)}{x} \tag{3.23}$$

where $M_{1.5}$ is the mass within $1.5\,h^{-1}$Mpc of the centre of the halo and $g(x)$ is given by Bartelmann (1996) as:

$$g(x) = \ln\frac{x}{2} + \frac{2}{x^2-1}\tan^{-1}\sqrt{\frac{x-1}{x+1}} \quad (x > 1) \tag{3.24}$$

$$g(x) = \ln\frac{x}{2} + \frac{2}{1-x^2}\tanh^{-1}\sqrt{\frac{1-x}{1+x}} \quad (x < 1) \tag{3.25}$$

$$g(x) = \ln\frac{x}{2} + 1 \quad (x = 1) \tag{3.26}$$

Combining Equations 3.10 and 3.26, we can deduce

$$\theta_q = \theta - \frac{D_{ls}}{D_s}\frac{4GM_{1.5}}{r_s c^2 g(1.5Mpc/r_s)}\frac{g(x)}{x} \tag{3.27}$$

$$\frac{d\theta_q}{dx} = \frac{d\theta}{dx} - \frac{D_{ls}}{D_s}\frac{4GM_{1.5}}{r_s c^2 g(1.5Mpc/r_s)}\frac{d\frac{g(x)}{x}}{dx} \tag{3.28}$$

and substitute in Equation 3.20 to ultimately find

$$\frac{\theta_q}{x} = \frac{r_s}{D_l} - \frac{D_{ls}}{D_s}\frac{4GM_{1.5}}{r_s c^2 g(1.5Mpc/r_s)}\frac{g(x)}{x^2} \tag{3.29}$$

$$\frac{d\theta_q}{dx} = \frac{r_s}{D_l} - \frac{D_{ls}}{D_s}\frac{4GM_{1.5}}{r_s c^2 g(1.5Mpc/r_s)}\frac{d\frac{g(x)}{x}}{dx} \tag{3.30}$$

which can be readily substituted into Equation 3.21 and numerically solved to derive the magnification. The magnification dictates the expected form of the correlation function as outlined in Equation 3.14.

### 3.4.3 Determining $\Omega_m$ in Groups

We will now briefly outline how the mass density of the Universe might be determined from our study, noting that the value thus derived is purely a measure of $\Omega_m$ in galaxy groups. $\Omega_m$ may be defined

$$\Omega_m = \frac{\rho_0}{\rho_{crit}} = \frac{8\pi G}{3H_0^2}\rho_0 \tag{3.31}$$

where $\rho_0$ can be estimated as the product of the space density of galaxy groups and the average mass of a dark matter halo. Equations 3.14 and 3.17 relate the average velocity dispersion of the SIS to the observed correlation function. The average mass of the SIS

may be derived from the average velocity dispersion, yielding an estimate of $\Omega_m$ in groups of

$$\Omega_m = \frac{8\pi \mathrm{G}}{3H_0^2} n M_{\mathrm{SIS}} = \frac{8\pi^2}{3H_0^2} n\sigma^2 r \tag{3.32}$$

where $n$ is the space density of galaxy groups and $r$ refers to the extent of the dark matter halo or, roughly, the extent of correlation. Similarly, we can express the correlation function in terms of the mass within $1.5\,h^{-1}$Mpc of the centre of an NFW halo. Theoretically, then, $\Omega_m$ is alternatively given by

$$\Omega_m = \frac{8\pi \mathrm{G}}{3H_0^2} n M_{\mathrm{NFW}} \geq \frac{8\pi \mathrm{G}}{3H_0^2} n M_{1.5} \tag{3.33}$$

where $M_{1.5}$ is the mass within $1.5\,h^{-1}$Mpc of the centre of an NFW halo.

## 3.5 Results

The models described in the previous section refer to a slightly different situation than hitherto discussed; the lensing of background flux around the *centres* of dark matter profiles, and so describe the correlation function of QSOs against the *centres* of galaxy groups. Previously, to compare our results with Boyle, Fong & Shanks (1988), we have measured the correlation between QSOs and individual galaxies in groups. To facilitate comparison with models, we now cross-correlate QSO positions with group centres, rather than group galaxies. Additionally, in prior sections of this chapter, larger groups of galaxies have been weighted more highly as each individual galaxy position within the group would be counted. We fit models to data where the correlation function is unweighted - each cluster is considered equally.

The cross-correlation between 2QZ QSOs and the centres of galaxy groups objectively determined from the APM South and the SDSS EDR is displayed in Fig. 3.13. For comparison, both the weighted and unweighted cross-correlations between QSOs and the centres of galaxy groups are shown. The weighted result (crosses) is a reflection of the analysis in Section 3.3.2, the QSOs are counted against each galaxy in a group, so larger groups are lent more significance. The unweighted result is not biased by the size of the groups. Binning the data in a sole bin out to 10 arcminutes and measuring the *rms* field-to-field variation in this bin, the average anti-correlation of the *weighted* result is $-0.049$

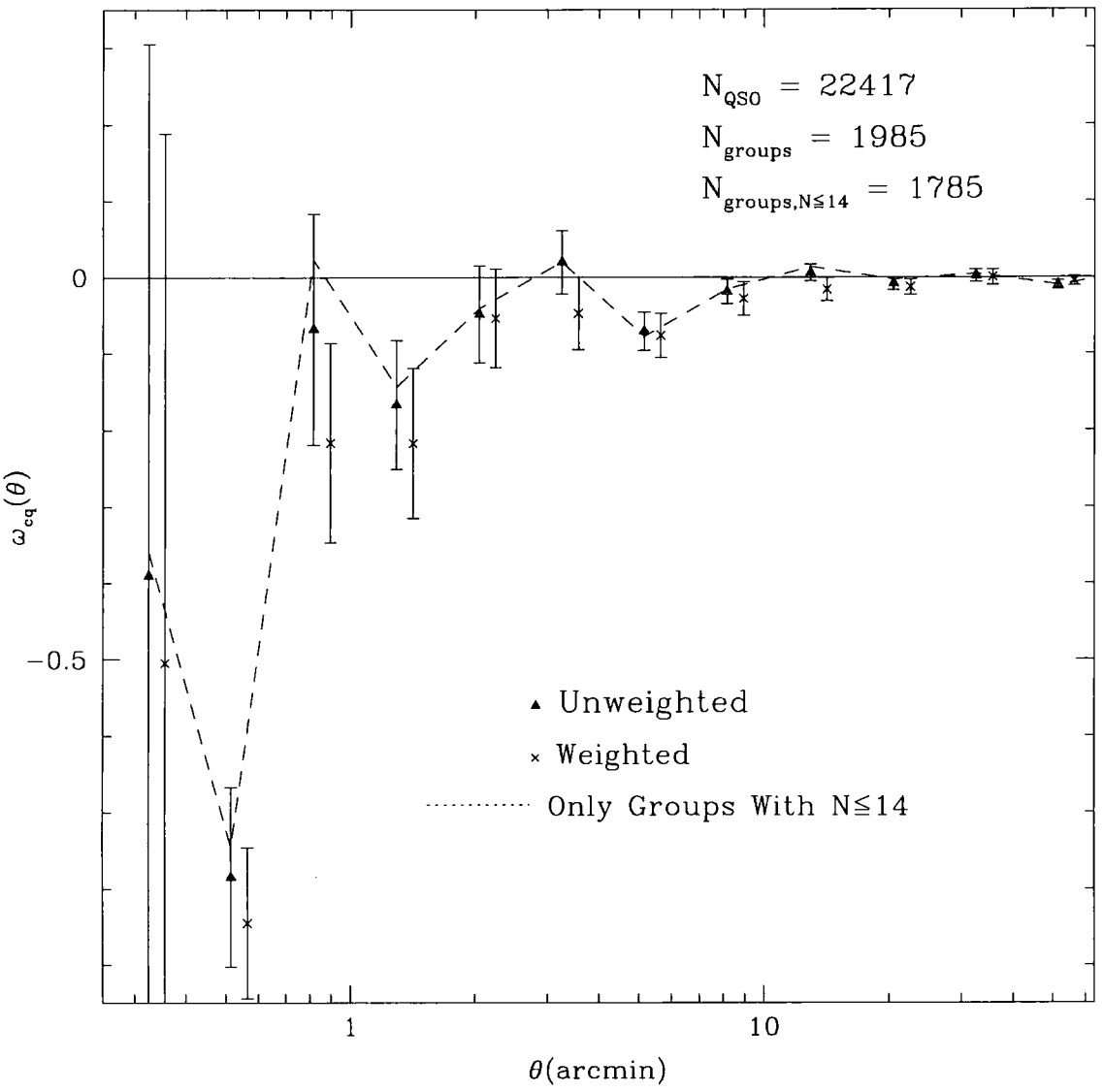


Figure 3.13: The cross-correlation between 2QZ QSOs and objectively derived galaxy groups of limiting magnitude $b_J = 20.5$ combined for both the southern and northern 2QZ strips. The numbers of each sample within the boundaries of the 2QZ are displayed. Crosses are derived by weighting each galaxy group by the number of galaxies in the group and have been offset slightly for display purposes. Triangles are derived counting each group equally. Errors are field-to-field. Also marked as a dashed line is the unweighted result for only those groups with 14 or fewer members. Errors on the dashed line are as plotted for the triangles.

with a significance of $2.9\sigma$. Comparing with the results in Section 3.3.2 , the strength and significance of the result weighted by galaxy number proves essentially identical whether we correlate QSOs with galaxies in groups or the group centres. The *unweighted* result has an average anti-correlation out to 10 arcminutes of $-0.034$ with a significance of $2.9\sigma$. When we do not weight the cross correlation by the number of galaxies in the group, the *strength* of the result thus drops by 30 per cent compared to the anti-correlation between QSOs and groups outlined in Section 3.3.2. We can deduce that the anti-correlation signal is stronger for larger groups, as would be expected if it is due to lensing.

To further demonstrate that the cross-correlation signal is not biased up by the handful of large clusters in our sample, we have cross-correlated QSOs against a sample of galaxy groups that has the largest 10 per cent of objects removed. The resulting sample of galaxy groups has between 7 and 14 galaxies per group. The anti-correlation for this sample is plotted as the dashed line in Fig. 3.13 and has essentially identical error bars to the full (unweighted) sample (marked with triangles). There is practically no difference in the result if we remove the largest 10 per cent of objects from our group sample, the anti-correlation signal thus derived has a strength of $-0.032$ and a significance of $2.5\sigma$. Note that even this weaker anti-correlation is outside the $2\sigma$ limit on dust in groups causing the anti-correlation, discussed in Section 3.3.3.

We model the lensing groups as dark matter haloes, either Singular Isothermal Spheres or NFW profiles. In the case of the SIS, the free parameter is the velocity dispersion of the sphere. The free parameter in the NFW model is the mass within $1.5\,h^{-1}$Mpc of the centre of the halo centre. For the lensing analysis, we use a faint-end QSO number-count slope of $\alpha = 0.29$ (see Chapter 2). The overwhelming majority of QSOs in the 2QZ lie fainter than the knee of a broken power-law model (only 10 per cent of 2QZ QSOs are brighter than $b_J = 19.1$). We have reproduced models both approximating the number-magnitude counts as a single power-law with slope $\alpha = 0.29$ and using the full smoothed power law determined in Section 2 (see Equation 2.2) and find no significant difference between the two approaches. We illustrate this point in Fig. 3.14, where we show the SIS lensing model described in Section 3.4 both assuming the approximation to the faint-end slope of the integrated QSO number-magnitude counts (see Equation 3.13) and a fit to the entire 2QZ (and 6QZ) integrated QSO number-magnitude counts (see the SPL model fitted to the integrated counts in Fig. 2.5). Assuming a single faint-end slope power law approximates the entire integrated QSO number-magnitude counts influences the modelled value of $\omega(\theta)$ by at most 0.025. Throughout the remainder of this chapter,

we adopt the approximation of Equation 2.2 but note that making this approximation has no influence on any of our analysis.

The lens and source distances required by the lensing models (see Equation 3.17, for example) are taken from the medians of the redshift distributions displayed in Fig. 2.6. The redshifts are translated to distances using an Einstein-de Sitter cosmology.

When fitting models, we use an estimate of the error based on 500 "mock" QSO catalogues of the same size and angular completeness as the 2QZ, calculating the cross-correlation between these mock catalogues and our galaxy group centres in the normal way, and then determining the *rms* errors between the catalogues. We fit a reasonable model to the anti-correlation result, then distribute the random placement of QSOs in the mock catalogues to reflect that model. So, if the cross-correlation fit has a value of -0.1 at 2 arcminutes, a random QSO that is generated 2 arcminutes from a group centre is only 90 per cent as likely to be included in the mock QSO catalogue as one in the field. Similarly, a QSO that then lies within 2 arcminutes of 2 group centres is 81 per cent as likely to be included. We have tested the independence of these errors by measuring the covariance of points in adjacent bins averaged over the 500 simulations, finding the covariance insignificant on all scales.

### 3.5.1 Model fits

Figure 3.15 shows the best-fit models for the SIS and NFW lensing haloes obtained by minimising the $\chi^2$ statistic. The errors are calculated from the standard deviation in the anti-correlation signal of 500 mock QSO catalogues as outlined above, with "mock" QSOs distributed according to a reasonable model. In the SIS case, mock QSOs are distributed according to a $\sigma = 600$ km s$^{-1}$ model when calculating errors. In the NFW case, mock QSOs are distributed according to a $M_{1.5} = 3 \times 10^{14}\ h^{-1}$ M$_\odot$ model when calculating errors. For brevity, the models are both displayed in Fig. 3.15 against the error bars calculated for the NFW case. We determine the best-fitting models out to 10 arcminutes, the extent of the anti-correlation. Once the best-fitting model is determined, the errors are scaled so the reduced $\chi^2$ is equal to 1 and then $1\sigma$ error-bars on the best-fitting model are calculated from this renormalised $\chi^2$ distribution. The best-fit SIS has a velocity dispersion of $\sigma = 1156 \pm^{93}_{327}$ km s$^{-1}$ with a reduced $\chi^2$ of 0.8. The best-fit NFW has a mass of $M_{1.5} = 1.2 \pm 0.9 \times 10^{15}\ h^{-1}$ M$_\odot$ with a reduced $\chi^2$ of 1.2. The data cannot distinguish variations in the NFW $\gamma$ parameter in the range $\frac{1}{3} < \gamma < 1$ (see Equation 3.22). As $\gamma$ increases, the predicted anti-correlation decreases below 1 arcminute and a test between

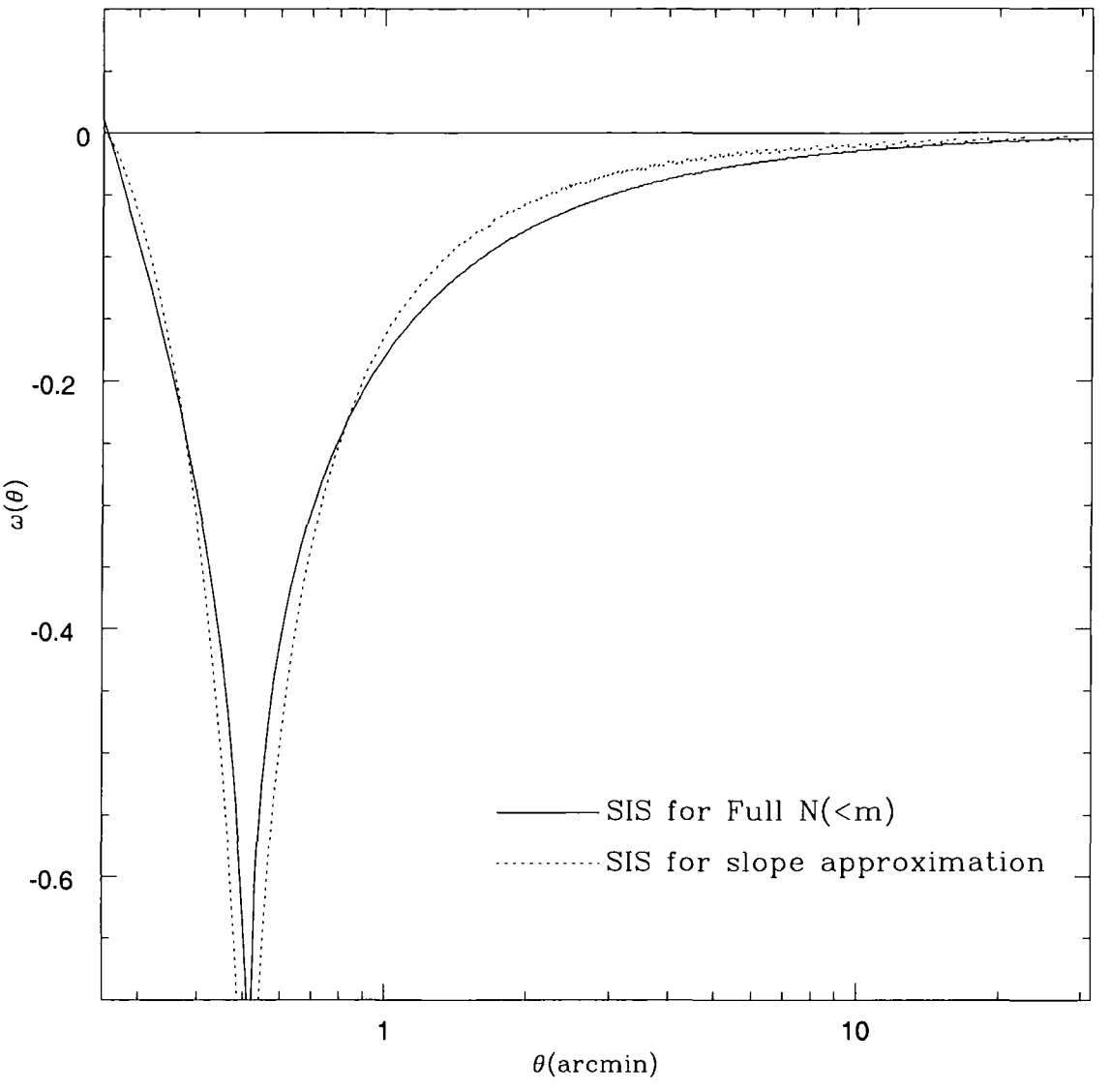


Figure 3.14: Equation 3.13 dictated the approximation made by many previous authors (Boyle, Fong & Shanks 1988, Narayan 1989, Croom & Shanks 1999) to the faint-end slope of the integrated QSO number-magnitude counts. In this figure, we plot our SIS lensing model (outlined in Section 3.4) both assuming this faint-end slope approximation and interpolating over the full form of the integrated QSO number-magnitude counts. The approximation effects the modelled value of $\omega(\theta)$ by (at most) 0.025 across the entire range plotted.

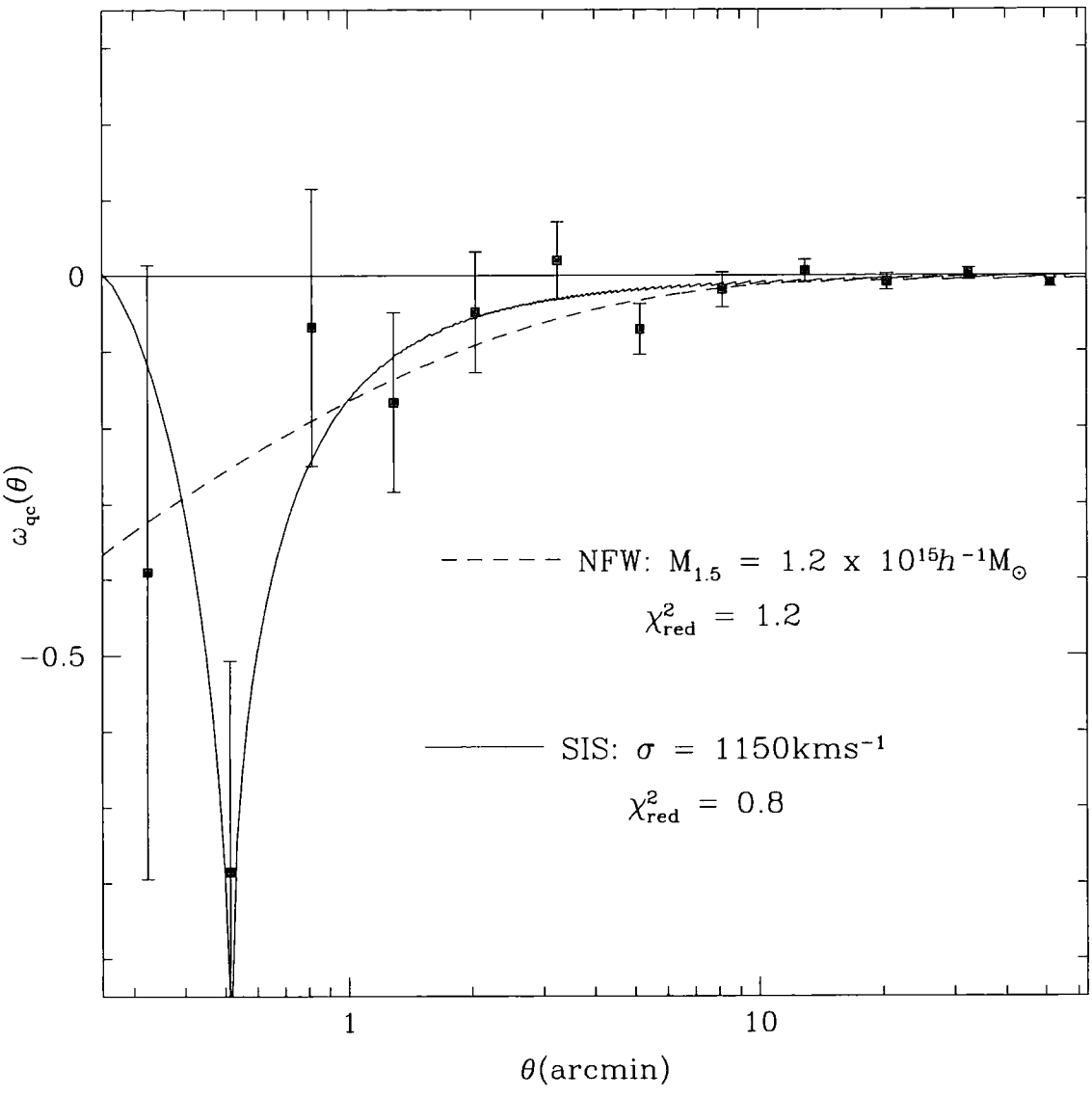


Figure 3.15: The cross-correlation between 2QZ QSOs and objectively derived galaxy groups of limiting magnitude $b_J = 20.5$ combined for both the southern and northern 2QZ strips together with best-fit models. Errors are standard deviations on the best-fitting NFW model derived from 500 mock QSO samples.

these cases should be possible in bigger datasets. In the current datasets there is hardly enough power to distinguish between the best-fitting SIS and NFW models.

The Einstein Radius of the best-fitting SIS model is around 30 arcsec, which results in a radical dip in the solid line in Fig. 3.15, corresponding to the terms in the denominator of Equation 3.17 being equal. At separations within the Einstein Radius, we enter the strong lensing regime and the SIS model predicts that each source QSO will produce two images (see Appendix C). One of these images has already been covered in the SIS model prediction for scales larger than the Einstein Radius. The second image appears within the Einstein Radius and might be considered a further prediction of the SIS model. The divergent nature of the SIS density profile on small scales means the dip in Fig. 3.15 may be predicted to occur at a scale which is unphysically large. Details of the model near the Einstein Radius do not affect the fitted SIS masses much, as evidenced by their similarity to those masses determined from fitting NFW profiles (for which the Einstein Radius is much smaller). The full consequences of strong lensing in the 2QZ, including an analysis of the numbers of multiply-lensed QSO systems, are discussed elsewhere (Miller et al. 2003).

As most of our model analysis is made in the weak lensing regime, we might be wary of any fit to the smallest scale data in the SIS case. As a consistency check, we have used mock QSO catalogues to make a direct test of the significance of rejection of $\sigma = 600$ km s$^{-1}$ and $\sigma = 1140$ km s$^{-1}$ SIS models, since the mock catalogues were produced for these specific cases. We determine how often cross-correlating mock QSO catalogues with galaxy groups can return an anti-correlation of significance $2.9\sigma$, as found for the real 2QZ data. We display this result in Fig. 3.16. We have created 250 mock QSO catalogues with the same size and completeness as the 2QZ. These are then cross-correlated against our galaxy groups and the strength of the cross-correlation is measured for each mock QSO catalogue. The measure of the significance of a cross-correlation is as we have used throughout this chapter (and will use throughtout this thesis), based on calculating the cross-correlation function and error in one bin out to 10 arcminutes. The mock QSO catalogues may also be distributed, as outlined above, to reflect various models for the lensing halo. We find that an anti-correlation of $2.9\sigma$ is measured 0/250 times if there are no lensing haloes, only 7/250 times ($2.2\sigma$) if the lensing haloes are modelled as an SIS with a 600 km s$^{-1}$ velocity dispersion (roughly equivalent to $\Omega_m = 0.3$) and 137/250 times for an SIS with an 1140 km s$^{-1}$ velocity dispersion (roughly equivalent to $\Omega_m = 1$). If we had chosen to plot the *strength* of the anti-correlation rather than

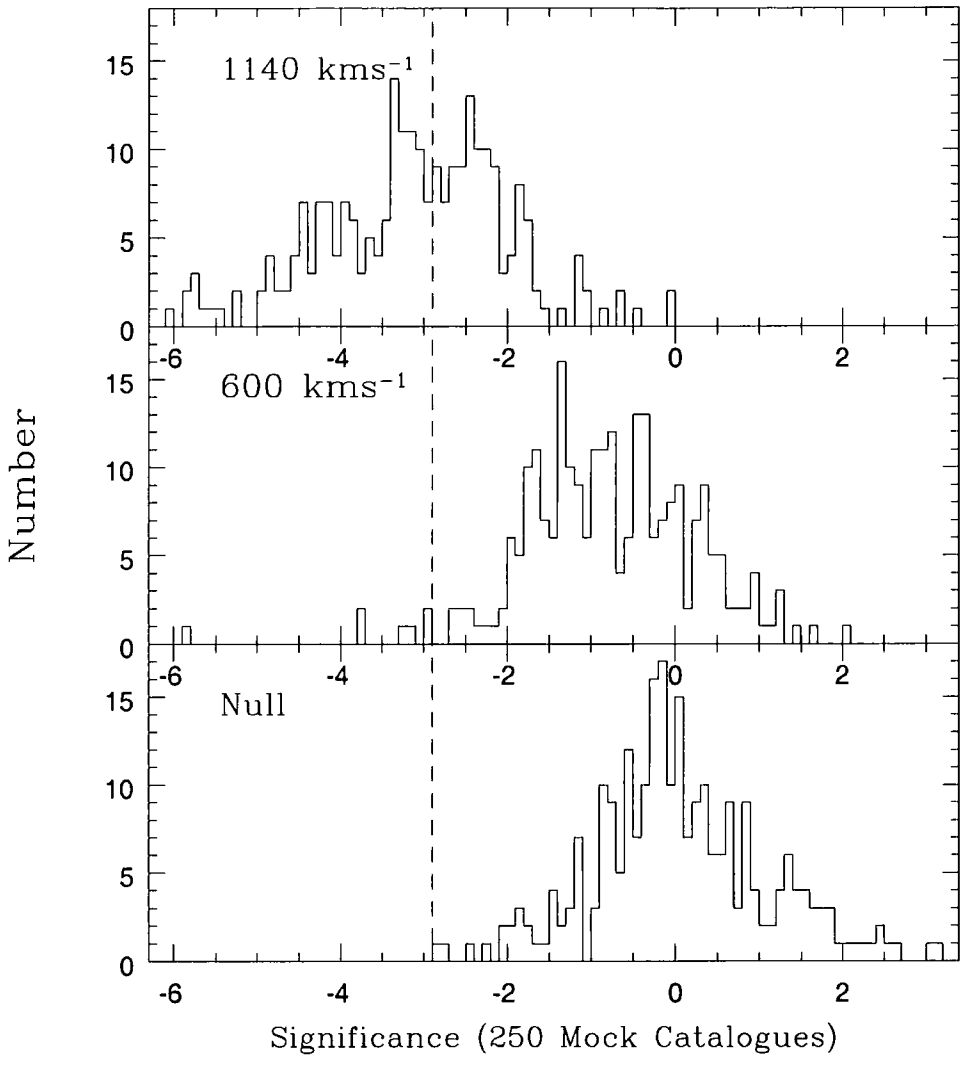


Figure 3.16: The distribution of the strength of correlation for cross-correlations between 250 mock QSO catalogues and objectively derived galaxy groups. The mock QSO catalogues have been weighted to reflect a number of different possible SIS halo models. The significance of the cross-correlation is measured out to 10 arcminutes for each mock QSO vs galaxy group cross-correlation function. Cross-correlating 2QZ QSOs against galaxy groups resulted in an anti-correlation of $-2.9\sigma$, which is marked as the dashed line.

the significance, we would find that a 600 km $s^{-1}$ model produces an anti-correlation of strength -0.034 (as found for the data) only 13/250 times ($1.9\sigma$). Either way, these results reject the 600 km $s^{-1}$ model at about the 5 per cent significance level. Note that this is a slightly stronger rejection than the error bars quoted in the first paragraph of this section, which correspond to a $1.7\sigma$ rejection. The discrepancy is likely because renormalising the reduced $\chi^2$ statistic to 1 is not an entirely fair representation of the error. If this renormalisation is not carried out, the rejection of the 600 km $s^{-1}$ model rises to $1.9\sigma$.

Wu et al. (1996) have pointed out that the inclusion of a uniform plane in modelling a dark matter halo may be considered a good reflection of the lensing influence of large-scale structure. Following the model of Croom & Shanks (1999), we have considered the effects of including such a plane in our SIS model. Including a plane of dark matter with our SIS has no influence on the best-fit model or its error-bars. In fact, the only effect worth remarking is to slightly lower our rejection of low velocity dispersion models. The best-fit SIS model rejects 600 km $s^{-1}$ at a $1.7\sigma$ level. When a plane is included, this rejection drops to $1.5\sigma$. If we consider the group centre auto-correlation function ($\omega_{cc}$), then we might expect the clustering of groups to have little effect on our signal. In Fig 3.17 we display $\omega_{cc}$, combined for all of our group data. A power-law fit of $\omega_{cc} \sim 0.1\theta^{-1}$ is a good match to the data. Integrating under $\omega_{cc}$ out to 10 arcminutes suggests that there might be as high as a 40 per cent contribution to our signal purely from the clustering of groups (i.e. integrating over Equation 3.4 to find the *excess* probability of one group being close to another).

There is still some debate over a number of the parameters used in the modelling process. Changing the QSO number-count faint-end slope to $\alpha = 0.34$ raises the SIS model estimate of the velocity dispersion by $\sim$ 10 per cent and lowering the index to $\alpha = 0.24$ lowers the estimate about 10 per cent. Changing to a cosmology with $\Omega_m = 0.3$ and $\Lambda = 0.7$, to model lens-source separations raises the estimate by less than 5 per cent. In the case of the NFW profile, lowering $\alpha$ to 0.24 lowers the mass estimate by $\sim$ 30 per cent, raising $\alpha$ to 0.34 raises the mass estimate by $\sim$ 50 per cent and changing to a Concordance cosmology raises the mass estimate by less than 5 per cent. In short, cosmology does not really affect our estimates but the exact gradient of the faint-end slope of the QSO number-magnitude counts may well be important, especially for the NFW profile. Reducing the QSO number-count faint-end slope to $\alpha = 0.15$, a very strong minimum value (being the differential counts slope determined in Section 2.3) would reduce the NFW mass estimate

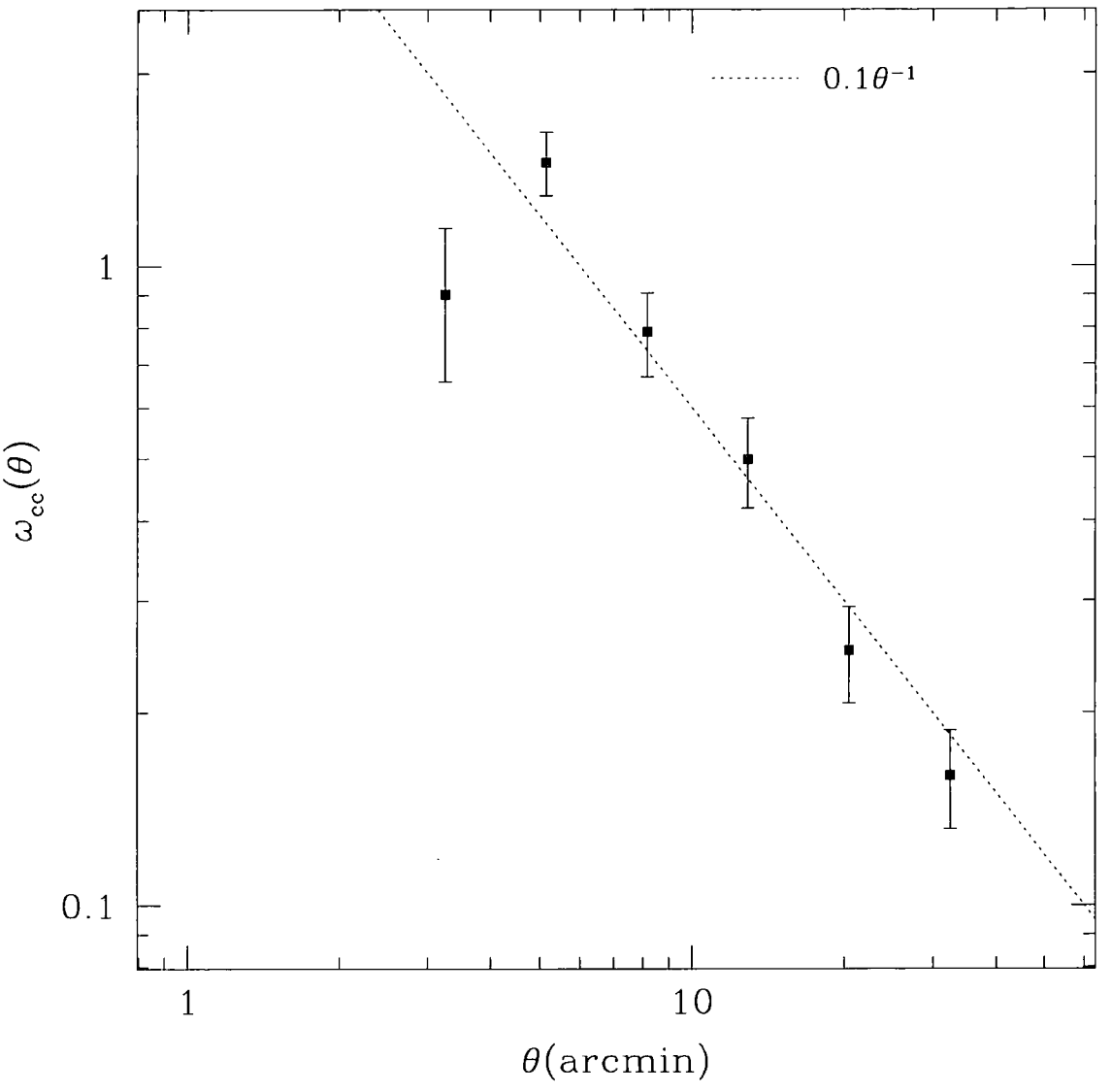


Figure 3.17: The auto-correlation function for groups with 7 or more members combined for our APM and SDSS galaxy group data. The data agree well with the group auto-correlation from Stevenson, Fong & Shanks (1988). An approximate fit to the data of $\omega_{cc} \sim 0.1\theta^{-1}$ is displayed. The slope of this fit is in good agreement with the Abell auto-correlation function of Abell clusters (Bahcall & Soneira 1983). Errors on the data are field-to-field.

by 70 per cent, bringing the NFW model mass prediction in line with a ΛCDM cosmology. It is possible that the magnification values generated by groups are large enough (Barber et al. 2000) to reach so faintly into the QSO counts that the value of $\alpha$ is as low as the differential counts value. This could conceivably allow a ΛCDM cosmology combined with an NFW profile to explain the anti-correlation between QSOs and galaxy groups.

### 3.5.2 Estimating $\Omega_m$ in Groups

In Section 3.4, we also note how estimates of the average group mass, or velocity dispersion, can be turned into estimates of the mass density of the Universe that is associated with groups (see Croom & Shanks 1999). The sky density of our groups is around $0.85\,\mathrm{deg}^{-2}$ in both the NGC and SGC strips. Note that this is significantly larger than the density of Abell clusters ($\sim 0.1\,\mathrm{deg}^{-2}$). Croom & Shanks (1999) have estimated the typical space density of such groups as $3 \pm 1 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$. Using this value, Equation 3.32 may be written

$$\Omega_m = \left( \frac{\sigma}{1125 \mathrm{kms}^{-1}} \right)^2 \tag{3.34}$$

assuming that the extent of the anti-correlation signal is $\theta = 10 \pm 2$ arcminutes ($r = 1 \pm 0.2\, h^{-1}\mathrm{Mpc}$). Equation 3.33 is equivalently

$$\Omega_m = \frac{M_{1.5}}{9.2 \times 10^{14} h^{-1} \mathrm{M}_\odot} \tag{3.35}$$

Using our best-fit estimates for the SIS returns a value of $\Omega_m = 1.06 \pm^{0.51}_{0.61}$ and for the NFW $\Omega_m \sim 1.3$, with large error.

The large error in Croom & Shanks (1999) value for the space density of groups remains a dominant systematic in our estimates of $\Omega_m$ and needs to be reassessed when group redshifts become available. Some additional error may be introduced by a lack of accurate redshift determinations for our galaxy groups. Groups of galaxies that are actually greatly separated in redshift may accidentally align and thus be counted as a single halo, although it is unclear to what extent this contamination could influence our lensing results, as any associations along the line of sight still trace an increase in the intervening mass distribution. Certainly, associations of galaxies that are separated greatly in redshift will not have dark matter profiles like the SIS and NFW profiles used in our modelling, being more like filaments than single haloes. We address these points

in Chapter 4. In the meantime, our estimates of $\Omega_m$ associated with groups continue to appear as high as those found by Croom & Shanks (1999) However, our errors have increased compared to the estimates of Croom & Shanks (1999), meaning the rejection of $\Omega_m \sim 0.3$ is only at the 1-2$\sigma$ level and we have noted that there may be further systematic errors still to be taken into account. Nevertheless, it is still tantalising that a $\sigma = 600$ km s$^{-1}$ dispersion is acceptable in at most 13/250 simulations for groups as numerous as those used here and this clearly motivates the application of this technique in larger QSO and galaxy group surveys.

### 3.5.3 Further Discussion

We have determined that the anti-correlation between galaxy groups and QSOs is best fit by a high-mass NFW model ($M_{1.5} \sim 1.2 \times 10^{15}\ h^{-1}$ M$_\odot$) or a high-velocity dispersion SIS model ($\sim$ 1150 km s$^{-1}$). Although the preference for the 1150 km s$^{-1}$ SIS group velocity dispersion is only at the $\sim 2\sigma$ level in the current data, taking this result together with the previous independently derived result of Boyle, Fong & Shanks (1988) and also with strong positive correlations seen in bright QSO samples (Williams & Irwin 1998, Gaztañaga 2003), it is clearly worth considering the implications if the amplitude of the anti-correlation was correct and caused wholly by weak lensing. Dynamical analysis of 2dF galaxy z-space distortions, results in an estimate of $\Omega_m^{0.6}/b \sim 0.43 \pm 0.07$ (Peacock et al. 2001) and so measurement of $\Omega_m \sim 1$ would imply $b \sim 2.5$. Although there is no immediate contradiction with the result presented a contradiction does arise when current CMB constraints on the mass power spectrum are included. These suggest that the galaxy power spectrum is approximately unbiased, implying $\Omega_m \sim 0.3$, in contradiction at the $2\sigma$ level with our best-fit result.

There are other constraints on the mass of galaxy groups which are in contradiction with our best-fit velocity dispersion. In particular, Hoekstra et al. (2001) have used shear to measure weak lensing of galaxies behind CNOC groups, finding $\sigma \sim 300$ km s$^{-1}$ and $M/L \sim 200h$M$_\odot$/L$_\odot$. These translate into a value of $\Omega_m = 0.19$. It is not clear why there is a difference between their results and ours. If the result of Hoekstra et al. (2001) is correct, we would have to appeal to moderate statistical fluctuations to explain our high anti-correlation amplitude in an $\Omega_m = 0.3$ model. Shear studies behind nearby galaxy groups which also have QSO lensing data might be valuable.

It should also be noted that the number of galaxies detected in N $\geq$ 7 groups corresponds to only 7.1 per cent of the total number of galaxies. In Fig. 3.18 we display

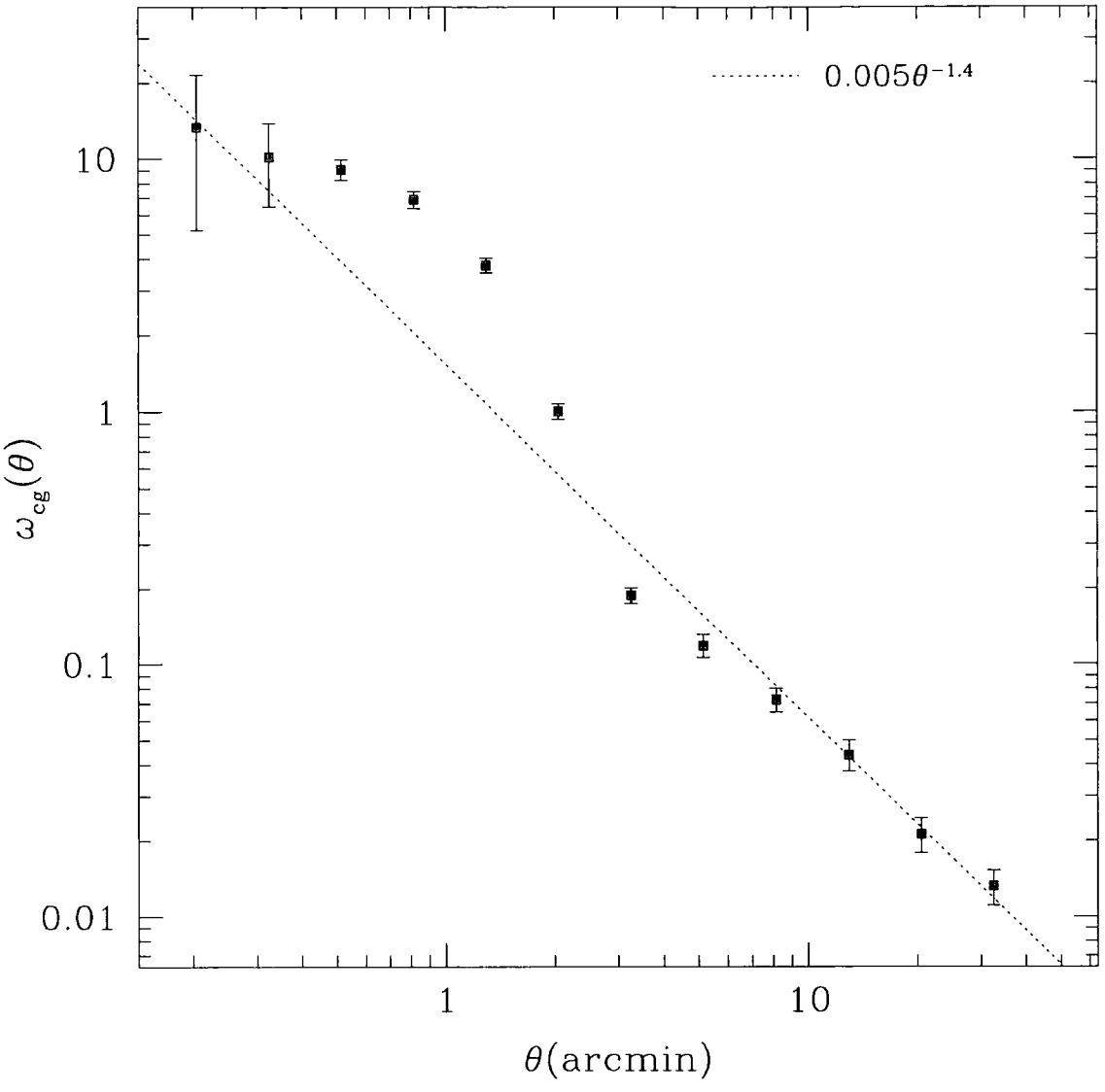


Figure 3.18: Groups with 7 or more members selected from the APM and SDSS are cross-correlated against all galaxies in the APM and SDSS that are contained within the boundaries of the 2QZ. The data are in reasonable agreement with $\omega_{cg}(\theta)$ for similarly selected groups from Stevenson, Fong & Shanks (1988). On scales larger than the typical group radius ($\theta > 2$ arcminutes), the data are reasonably fit by a power law with a slope of $-1.4$, in good agreement with Seldner & Peebles (1977). Errors on the data are field-to-field.

$\omega_{cg}$, the cross-correlation of the centres of groups (with more than 7 members) against all galaxies in the APM or SDSS (that are contained within the boundaries of the 2QZ). Stevenson, Fong & Shanks (1988) note that the kink in $\omega_{cg}$ around the scale of the typical group radius of $\sim$ 2 arcminutes (see Equation 3.3) is an artefact of the group-detection procedure. On all other scales, $\omega_{cg}$ is fit well by a model with a power-law slope of -1.4 (Seldner & Peebles 1977). Using Equation 3.4, and integrating $\omega_{cg}$ out to a radius of $\theta < 2'$ in comparison to $2' < \theta < 10'$ (the extent of our anti-correlation signal), suggests that the total number of galaxies associated with our groups is actually only $\sim$ 40 per cent. Meaning that, assuming the mass properties of galaxies are similar to that of groups, our $\Omega_m$ estimate could be as high as $\sim 2.5$.

## 3.6 Conclusions

We have sought the effects of weak gravitational lensing in correlations between 2QZ QSOs and galaxy groups derived from the SDSS or from the Southern APM. We confirm that there is a distinct ($\sim 3\sigma$) anti-correlation between objectively determined galaxy groups and QSOs (Boyle, Fong & Shanks 1988). The anti-correlation is fit well by supposing its cause is gravitational lensing through dark matter haloes, either NFW haloes or Singular Isothermal Spheres, but requires more mass than models with $\Omega_m = 0.3$ would suggest, at the $\sim$ 1-2$\sigma$ level. Larger QSO samples could not only better distinguish the amplitudes of the anti-correlation predicted by cosmological models but may also be able to distinguish between different forms for dark matter halo profiles.

The observed anti-correlation does not appear to be caused by a selection effect due to the limited spacing of 2dF fibres. We also rule out the idea that intervening dust causes the signal (Boyle, Fong & Shanks 1988). Our 95 per cent upper limits on reddening in $b_J - r$ are 0.012, which corresponds to $A_B < 0.04$ mag assuming the usual Galactic dust law, which is in good agreement with Nollenberg, Williams & Maddox (2003). To explain the anti-correlation by dust would need $A_B \approx 0.2$ mag (see Boyle, Fong & Shanks 1988); the dust hypothesis could then only be saved by appealing to either a smooth distribution of grey dust or a patchy distribution of heavily line-of-sight distributed dust around galaxy groups. It is also not straightforward to rule out the hypothesis that *both* dust and lensing play some combined role in producing the anti-correlation signal.

It seems that weak lensing remains the likely explanation for the anti-correlation between QSOs and galaxy groups. The strength of the anti-correlation suggests that there

may be more mass hidden in galaxy groups than many previous estimates require, making further study of this phenomenon worthwhile. The models described in Section 3.4 to describe the form of the correlation function remain simplistic averages and it would be worth performing some large, high-resolution simulations of the expected lensing influence of foreground mass on QSO distributions in different cosmologies. Accurate ray tracing through N-body simulations would predict the expected anti-correlation for different cosmologies. Additionally, it is probable that a large proportion of our groups are really chance alignments of galaxies that are actually greatly separated in redshift. In the next chapter, we discuss running our group detection procedure using galaxies from a large simulation, allowing us to determine how many of our groups are genuine haloes and how many are chance alignments.

On the observational side, there is impetus for a new faint QSO survey to determine the faint-end slope of the QSO number-magnitude relation. We discuss such a survey in Chapter 6. Redshift information from both mock catalogues and in the 2dF Galaxy Redshift Survey should be helpful in improving both the nature of our group sample and estimates of the lensing mass. The next chapter focuses on the cross-correlation of 2QZ QSOs against 2dFGRS galaxy groups (Eke et al. 2003). The 2dFGRS groups are selected to reflect the expected mass of underlying dark matter haloes, allowing us to compare the lensing mass inferred from QSO-group cross-correlations against the expected halo mass from simulations.

# Chapter 4

# *The Correlation of Faint QSOs and Dark Matter Haloes*

## 4.1 Introduction

In the previous chapter, we showed there is a significant anti-correlation between faint QSOs and groups selected as over-densities in the angular distribution of galaxies. When these galaxy groups are modelled as dark matter haloes, we find that the mass required to explain the anti-correlation is more than a ΛCDM model can account for. Though the veracity of the anti-correlation as a physical effect has been checked extensively, an obvious criticism of the modelling process may be raised. Namely, though the galaxy groups themselves are angular, two-dimensional projections, the haloes used to model them describe three-dimensional agglomerations of mass. A large number of galaxies that are aligned from our viewpoint may still cause a mass over-density that leads to a significant lensing effect but these galaxies may well be spread out in redshift space, meaning that a single halo may not be an appropriate model to describe this lensing.

It was rapidly realised that the objective algorithm proposed by Turner & Gott (1976), which we used to define groups in the previous chapter, picked out a variety of structures that often didn't fairly represent a physically associated set of galaxies. Occasionally, the algorithm doesn't even reflect what would be selected by an attentive observer - for instance, Kirshner (1977), found a torus of galaxies selected in the original analysis of Turner & Gott (1976) that enclosed a single galaxy, which was *not* assigned membership of the torus that surrounded it! More importantly, galaxies that are separated in redshift are frequently projected into a single group, if the main criteria for grouping the galaxies is their projected over-density in the plane of the sky (Reblinsky & Bartelmann 1999, White & Kochanek 2001). As galaxy redshift surveys have become more prevalent, more sophis-

ticated grouping algorithms have emerged that incorporate the redshift coordinate as an extra determinant in realising truly three-dimensional groups. These grouping algorithms mainly fall into two classes "percolation" (or "friends-of-friends") algorithms (Huchra & Geller 2002, Geller & Huchra 1983) and "hierarchical" algorithms Tully (1987). Loosely, both types of algorithm work by arriving at some appropriate "linking length" (which might be a function of many variables, particularly redshift) that divines when a galaxy should be considered associated with other galaxies. Percolation techniques then define all galaxies that are linked as a group. Hierarchical techniques, on the other hand, find structures that minimise the linking length, redefine the position of these larger structures, then once again minimise the linking length between these new structures.....and so on, until all galaxies have a place in a hierarchy from which a particular level is chosen to represent groups. Tully (1987) used this hierarchical technique to find 103 groups of at least 3 members from a sample of $\sim$ 2000 galaxies, Gourgoulhon, Chamaraux & Fouque (1992) used a similar hierarchical algorithm to define 264 groups of at least 3 members in a sample of $\sim$ 4000 galaxies, Garcia (1993) found 485 groups in a sample of $\sim$ 6000 galaxies, utilising both percolation and hierarchical algorithms. Redshift surveys have grown rapidly in recent years and a large number of objectively-defined three-dimensional groups have been sampled. Carlberg et al. (2001) identified 200 groups at intermediate redshift ($0.1 < z < 0.55$) from a sample of $\sim$ 3000 galaxies with redshifts, taken from the CNOC2 Survey. Percolation algorithms applied to $\sim$ 20,000 galaxies taken from the Updated Zwicky Catalog and Southern Sky Redshift Survey have defined a sample of nearly 1000 groups (Merchãn, Maia & Lambas 2000, Ramella et al. 2002). Most recently, the first 60,000 galaxies released from the 2dF Galaxy Redshift Survey (Colless et al. 2001) were probed with a percolation technique, resulting in the definition of over 2000 groups in total (Merchãn & Zandivarez 2002). The final 2dF Galaxy Redshift Survey (henceforth 2dFGRS) catalogue (Colless et al. 2003) represents the largest extant galaxy redshift survey, containing $\sim$ 220,000 galaxies with redshifts, targeted to a magnitude limit of $b_J = 19.45$. Eke et al. (2003) have recently used a percolation technique to define groups in this sample and find $\sim$ 13,000 groups with at least three members, a sample of the order of 10 times larger than the next largest objectively-defined redshift-space group catalogue. Conveniently, for our purposes, 55 per cent of the groups defined by Eke et al. (2003) lie within the confines of the 2dF QSO Redshift survey (henceforth 2QZ), allowing us to probe the statistics of QSO-group associations in these samples.

Groups taken from the catalogue of Eke et al. (2003) are particularly complementary

to our lensing-by-haloes modelling process outlined in Chapter 3, as Eke et al. (2003) select groups using a linking length that picks out structures reflective of dark matter haloes in N-body simulations. To do this, they first construct a *mock catalogue* that mimics the 2dFGRS, test a range of linking lengths on this mock catalogue and thus derive a linking length that groups galaxies that best reflect underlying haloes.

The construction of a mock catalogue is basically a two-step process (although, see Scoccimarro & Sheth 2002). First, an N-body simulation (see Jenkins et al. 1998 for a review of such simulations) is carried out to create a catalogue of dark matter particle positions. This part of the process contains information on how matter in the universe clusters under the effects of gravity. It is then necessary to decide which of the collection of dark matter particles "becomes" a luminous galaxy, which requires assumptions about how baryons behave within agglomerations of dark matter. On the simplest level, one may assume a physically motivated probability function (dependent on the density contrast of dark matter particles on a given scale compared to the average density of dark matter particles) and select galaxies from the catalogue of dark matter particles according to that function (Cole et al. 1998, Colley et al 2000). Alternatively, it is possible, under a number of ("semi-analytic") simplifying assumptions, to physically evolve the baryonic component of the universe in check with the evolution of the underlying dark matter and examine how galaxies emerge (Cole et al. 2000). As the physics of star-formation are still poorly understood, models of how galaxies form are indubitably somewhat inaccurate but mock catalogues *have* been constructed using semi-analytic galaxy formation (Diaferio et al. 1999). Independent of the model used to populate haloes with galaxies, the distribution of galaxies derived is ultimately tweaked so that the final mock catalogue shares the boundaries, auto-correlation function and radial selection function of the survey it mimics.

The purpose of this chapter is twofold. In Section 4.2, we use a mock catalogue (Cole et al. 1998) to study how many groups of galaxies determined by our angular group-finder are associated in redshift-space. In Section 4.3, we cross-correlate our QSO sample against a catalogue of groups determined in three-dimensions (Eke et al. 2003) from the 2dF Galaxy Redshift Survey (which are more appropriately modelled as a collection of haloes than the groups that we have determined as angular over-densities in the previous chapter) and fit the results using the NFW lensing model derived in Section 3.4 of Chapter 3. Section 4.4 draws together the information in this chapter and considers the implications when combined with the results of the previous chapter.

## 4.2 2D and 3D Group Statistics

It is worth considering what proportion of the groups defined in Chapter 3 are associated in redshift-space as, ultimately, this chapter will consider the cross-correlation between QSOs and a sample of groups selected as associations in redshift-space. In this section, we use a mock catalogue to investigate how many of the groups determined in Section 3.2 of Chapter 3 as angular over-densities on the plane of the sky are likely to be associated structures when considered in redshift-space. In this chapter, for brevity, we shall often refer to groups determined in the plane of the sky as "2D" groups and those determined using both angular information and redshifts, or defined in theoretical simulations using proper spatial coordinates as "3D" groups.

The "reality" of a 3D group or cluster in redshift-space is a more complicated concept than it seems. "Real" or "associated" structure is generally thought of in terms of spatial position - people grouping around a steet-performer or bees clustering around a hive - but in cosmology, coordinates are never entirely spatial. The redshift coordinate is a degenerate measure of both the Hubble Flow (which can be used to trace the spatial coordinate) and the peculiar velocity of objects within the Hubble Flow. Generally, then, a 3D group based a posteriori on the redshift coordinate is defined by a boundary in velocity space, or a difference in comoving coordinates. However, in theoretical simulations, where all of the spatial coordinates are know a priori, a group can be defined as an association of galaxies in proper space coordinates. This can lead to confusion about what is meant when referring to a 3D "group" of galaxies. Further, groupings of *mass* can be selected in theoretical simulations on the basis of density profiles, or how the mass drops with radius in proper spatial coordinates. To be distinctive, we will refer to groups selected "observationally" on the basis of their redshift coordinate as "groups in redshift-space" or "associated in redshift space". We will refer to galaxies that are grouped in proper spatial coordinates as "proper groups" or "associated in proper space". We will refer to groups selected from simulations on the basis of the underlying mass as "haloes" or "dark matter haloes".

### 4.2.1 The Mock Catalogue

To best investigate how many 2D groups determined in Section 3.2 of Chapter 3 persist in a 3D sample would require redshifts for each of the (large number) of galaxies that were assigned group membership. These galaxies were taken from the Sloan Digital Sky

Survey (henceforth SDSS) Early Data Release (henceforth EDR) and the southern APM Survey. Neither of these galaxy surveys currently has redshift information to $b_J \sim 20.5$, and no extant survey will provide redshifts to this depth over the southern APM and SDSS EDR regions (the main SDSS survey - of the northern sky to galactic latitude $b > 30$ degrees - will only obtain redshifts to $B < 19$). Our strategy will therefore be to consider a simulated sample of these galaxies, with full redshift information.

Simulating either the APM or SDSS to $b_J < 20.5$ with any degree of sophistication is by no means straightforward. Fortunately, mock catalogues of the APM Survey to this depth already exist. Shaun Cole (of Durham University) has produced catalogues that mimic the APM based on the procedures outlined in Cole et al. (1998) and using dark matter particles drawn from the large (1 billion particle) "*Hubble Volume*" simulations of the Virgo Consortium (see, for instance, Jenkins et al. 1998). These mock APM catalogues to $b_J < 20.5$ are currently available at `http://star-www.dur.ac.uk/~cole/mocks/hubble.html`.

The *Hubble Volume* simulations are described at some length in Evrard et al. (2002). They are the most voluminous simulations of the evolution of dark matter in our universe to date. N-body simulations essentially take a set of particles (of given mass) imposed with an initial spectrum of density fluctuations and evolve them over time under the influence of gravity. The gravitational force between particles usually takes the form of a Plummer Law ($F \propto [r^2 + \epsilon^2]^{-1}$) with a softening scale ($\epsilon$) included to prevent close particles suffering inordinately large interaction. The *Hubble Volume* simulations are constrained by the amplitude and shape of the initial density fluctuations - characterised by the power spectrum $P(k)$, the baryonic content of the universe (which modifies the shape of the initial density fluctuations) and the cosmology of the model universe, which determines the expansion rate of (and distances in) that universe. The amplitude of the initial mass fluctuations can be constrained using the parameter $\sigma_8$, the root-mean-square of density fluctuations in spheres of $8\,h^{-1}$Mpc radius at the present day. The shape of the initial power spectrum is defined $\Gamma = \Omega_m h$. The baryonic content of the universe, characterised by $\Omega_b$, the fraction of the critical density in baryons, modifies the shape of $P(k)$ according to the ansatz $\Gamma \simeq \Omega_m h \exp\left(-\Omega_b \left[\sqrt{2h} + \Omega_m\right]/\Omega_m\right)$ (Sugiyama 1995). Cosmology is codified in Hubble's Constant, $H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, the mass density of the universe compared to the critical density, $\Omega_m$, and the fraction of the closure density supplied by a cosmological constant $\Omega_\Lambda$. Constraints on these parameters are discussed in Chapter 1, the introduction of this thesis. We note that all of these parameters are

now well constrained by the power spectrum of WMAP data (Spergel et al. 2003). So, in simple terms, an initial set of particles of certain mass, distributed with a set of density fluctuations of shape $\Gamma$ is modified by $\Omega_b$, evolved under the influence of gravity in a cosmology defined by $\Omega_m$, $\Omega_\Lambda$ and $H_0$ and the amplitude of the resulting matter fluctuations at the present day are normalised to $\sigma_8$. The outputted set of mass particles comprises a simulation of the distribution of dark matter in the universe.

The mock catalogue we will use selects galaxies from the dark matter particles of the *Hubble Volume* using the probability function

$$P(\nu) = \begin{cases} \exp(\alpha\nu + \beta\nu^{3/2}) & \text{if } \nu \geq 0 \\ \exp(\alpha\nu) & \text{otherwise,} \end{cases} \tag{4.1}$$

where the probability is a function of $\nu = \delta/\langle\delta^2\rangle$, which is in turn a function of the density contrast, $\delta(\vec{r})$, defined $\delta(\vec{r}) = (\rho(\vec{r}) - \bar{\rho})/\bar{\rho}$, where $\rho$ denotes the density of dark matter particles. The free parameters, $\alpha$ and $\beta$ are then adjusted so that the auto-correlation function of the sampled galaxy distribution agrees with that measured for the APM Survey (Baugh, Cole & Frenk 1996). A final sample of galaxies is then chosen according to the radial selection function of the APM Survey and assigned magnitudes according to the luminosity function of the APM Survey (Loveday et al. 1992b).

The only available APM mock catalogues to depth $b_J < 20.5$ are based on a *Hubble Volume* simulation carried out in a ΛCDM cosmology, with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\sigma_8 = 0.9$, $\Gamma = \Omega_m h = 0.21$, $\Omega_b = 0.04$ and $H_0 = 70\,\mathrm{km\,s^{-1}}$ Mpc. The exact choice of cosmology does not ostensibly affect the properties of groups, though (Diaferio et al. 1999), so the cosmology of the underlying N-body simulation used to create the APM mock catalogue is not too important for our purposes. The resolution of the *Hubble Volume* simulations (i.e., the gravitational force softening scale, called $\epsilon$, above) is $0.1\,h^{-1}$Mpc. For use in the rest of this section, we select a ΛCDM APM mock catalogue from the 22 made publicly available by Shaun Cole at `http://star-www.dur.ac.uk /~cole/mocks/hubble_volume_lcdm03/`, and use the mock galaxies from it that fall within the confines of the 2dF Southern Galactic Cap (henceforth SGC) strip.

### 4.2.2 A 3D Grouping Algorithm

In Section 3.2, we described the algorithm, invented by Turner & Gott (1976), that we use to select 2D groups from the SDSS EDR and APM galaxy catalogues in the Northern and Southern 2QZ strips. Essentially, the Turner and Gott algorithm (henceforth "2D

algorithm") assigns a radius to each galaxy in the sample dependent on the sky density of its neighbouring galaxies compared to eight times the mean sky density, eight times the background being the level of over-density suggested by Turner & Gott (1976) to pick uncontaminated groups out of the field. We call this assigned radius the *isocontour* of each galaxy. Galaxies with overlapping isocontours are called groups. Naively, one might expect that this 2D grouping algorithm could be extended into redshift-space, so that we assign spheres to each galaxy depending on the space density of its neighbours (an "isosphere", if you will). In a static universe where the redshift coordinate garnered a precise measurement of distance, this would be an acceptable method. However, the expansion of our Universe is not the only phenomenon that can affect the wavelength of light from distant objects. One influence on galaxy redshifts are the peculiar velocites of the galaxies. Galaxies that have an intrinsic line-of-sight velocity towards our own Galaxy will have light that is blue-shifted relative to a stationary galaxy at the same cosmological distance. The most obvious example of this is the blue-shifted Andromeda galaxy, gravitationally attracted to the Milky Way and rapidly falling towards us - see, e.g., Rubin & D'Odorico (1969). Similarly, the light from galaxies that have a line-of-sight velocity away from us will be red-shifted relative to light from a stationary galaxy at the same cosmological distance. Thus, as galaxies orbit a central gravitational potential, such as in a group of galaxies, they appear to spread out in redshift-space - the so-called *fingers-of-God* effect.

When defining groups from the APM mock catalogue in the 2QZ SGC region using the 2D algorithm, it becomes apparent that the peculiar velocities of galaxies in groups will generally be larger than the comoving separation of galaxies, based on the angular isocontours the 2D algorithm assigns each galaxy. The largest isocontour (in terms of comoving distance rather than angular separation) assigned to a galaxy in the APM mock catalogue by the 2D algorithm is about 3 arcminutes, assigned to a galaxy at $z \sim 0.36$. For a ΛCDM cosmology, this corresponds to a separation (projected in the plane of the sky) of about $0.6\,h^{-1}$Mpc, equivalent to a difference in velocity of about $60\,\mathrm{km\,s^{-1}}$. Strauss, Ostriker & Cen (1998) measured the relative velocity of galaxies in high density regions (6 times the mean sky density) of the Optical Redshift Survey (Santiago et al. 1995) and found a typical relative velocity of about $800\,\mathrm{km\,s^{-1}}$ on scales of $1\ h^{-1}$Mpc. This result is in good agreement with ΛCDM simulations. Sheth & Diaferio (2001) studied the peculiar velocites of galaxies in the high resolution (softening length $\sim 30\,h^{-1}$kpc) GIF N-body simulations (Kauffmann et al. 1999), finding, for a ΛCDM

cosmology, an *rms* velocity of $700\,\mathrm{km\,s^{-1}}$ when considering the densest 5 per cent of dark matter particles in their simulations (i.e. on average, a *relative peculiar velocity* of $1000\,\mathrm{km\,s^{-1}}$). Note that the 2D groups determined in the previous chapter with 7 members or more contain about 10 per cent of the total number of galaxies. Thus the maximum physical separation of $60\,\mathrm{km\,s^{-1}}$ found from the assigned 2D isocontours in the APM mock catalogue is relatively small compared to the expected peculiar velocities of the galaxies. In fact, $60\,\mathrm{km\,s^{-1}}$ is actually smaller than the typical measurement errors on galaxy redshifts in the 2dFGRS (Colless et al. 2001)! The peculiar velocities of galaxies, rather than their physical separation, should therefore be the determining factor in finding groups in redshift-space.

Having determined which galaxies are in 2D groups using our 2D algorithm, we determine which of these groups persist when redshifts are known using a very simple percolation algorithm. If the separation in redshift-space of galaxies that are included in a 2D group is less than the maximum allowed peculiar velocity of a galaxy in that group, we assign it membership of the equivalent 3D group. Thus, our criterion for a galaxy that is already included in a 2D group to be included in a 3D group is

$$\frac{c\Delta z}{1+z} \leq \Delta v \tag{4.2}$$

where $c$ is the speed of light, $z$ is the redshift of the particular galaxy considered, $\Delta z$ is the difference in redshift between the particular galaxy considered and a neighbouring galaxy in the group (the group having being previously defined using the 2D algorithm) and $\Delta v$ is the maximum allowed relative velocity of a group. The criterion of Equation 4.2 thus defines a sort of "linking length" in redshift, beyond which galaxies will no longer be considered grouped.

Strauss, Ostriker & Cen (1998) found the typical relative velocities for galaxies in regions 6 times denser than the mean density to be $800\,\mathrm{km\,s^{-1}}$. Given that we are considering denser regions and on smaller scales, given that ΛCDM simulations suggest slightly higher relative velocities between galaxies ($\sim 1000\,\mathrm{km\,s^{-1}}$ relative velocity, or $700\,\mathrm{km\,s^{-1}}$ *rms*, in quite dense regions), given that results in the previous chapter suggests more mass on small scales than ΛCDM , and given that our goal is to retain as many proper 3D groups as possible from 3D groups defined in redshift-space, we nominally work with the criterion $\Delta v = 2000\,\mathrm{km\,s^{-1}}$, corresponding to twice the *rms* variations for the densest 5 per cent of particles in the ΛCDM simulations of Sheth & Diaferio (2001). We will also consider $\Delta v = 4000\,\mathrm{km\,s^{-1}}$ and $\Delta v = 1000\,\mathrm{km\,s^{-1}}$. Note that having a single $\Delta v$

for each galaxy in a group is a simplification, in reality, the maximal peculiar velocity for a group galaxy will be a function of both redshift and the mass of the group (Sheth & Diaferio 2001). However, assigning galaxies group membership is an extremely sophisticated play-off between contaminating the group with interlopers and missing group members (Eke et al. 2003), and so we will only consider 3D groups to first order. We will often refer to groups that have been defined (from a previously determined 2D group sample) using the criterion of Equation 4.2 as groups defined using *the 3D algorithm.* In the rest of this chapter, if we refer to "3D groups" rather than "proper 3D groups", or "3D groups in proper space" then we mean groups selected using the 3D algorithm. We will sometimes still use the general "groups selected in redshift-space" when we mean groups selected using a generic redshift-based selection, rather than specifically using Equation 4.2.

### 4.2.3 2D and 3D Group Comparison

In Fig. 4.1 we display the multiplicity function of groups determined from our galaxy samples - taken from the SDSS EDR in the region of the 2QZ NGC strip and from the APM survey in the region of the 2QZ SGC strip - against the multiplicity function for groups drawn from the APM mock catalogue. The multiplicity function, $f(N)$ is the fraction of the total number of galaxies that appear in groups with $N$ members. In Fig. 4.1, we plot 3D groups for $\Delta v = 2000\,\mathrm{km s^{-1}}$, where $\Delta v$ sets the redshift-space criterion in Equation 4.2. Groups determined from the mock catalogue using the 2D algorithm are in very good agreement with 2D groups from the observational data. However, as might be expected, groups determined from the mock catalogue using the 3D algorithm have a very different multiplicity from those determined using the 2D algorithm. The 2D algorithm places 49 per cent of galaxies in the field (i.e. not being associated with any other galaxy), where the 3D algorithm determines a 74 per cent fraction of field galaxies. The 2D algorithm finds 10 per cent of the galaxies lie in groups with 7 or more members ($N \geq 7$) but the 3D algorithm only assigns 5 per cent of galaxies to groups with $N \geq 7$. This trend holds for larger groups, for instance the 2D algorithm finds 4 per cent of the galaxies lie in $N \geq 15$ groups but the 3D algorithm only assigns 2 per cent of galaxies to groups with $N \geq 15$.

To further investigate how 2D groups with at least 7 members ($N \geq 7$) trace 3D associations in redshift-space, we define the following statistics

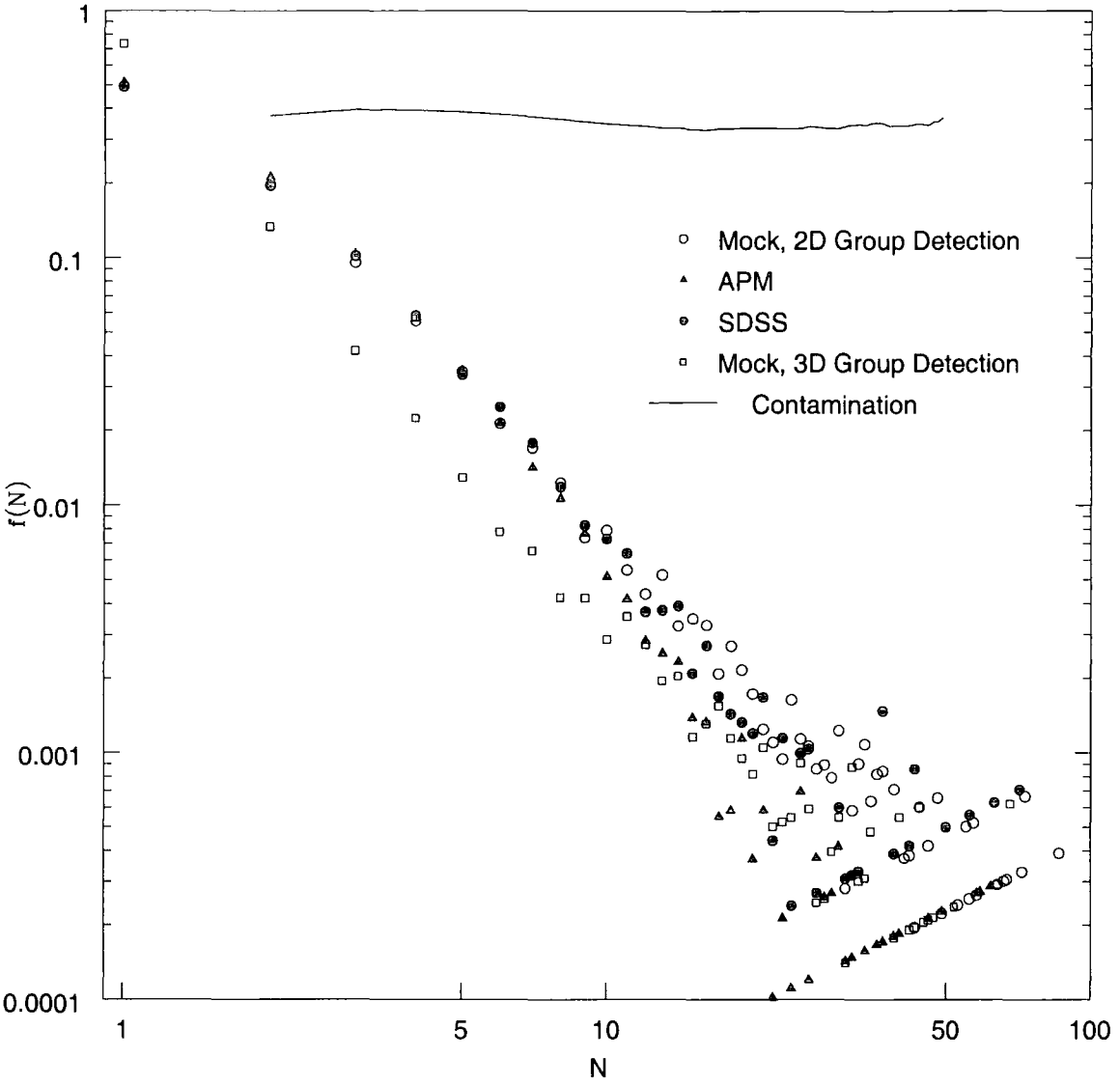


Figure 4.1: For the points, $f(N)$ denotes the multiplicity function - the fraction of galaxies that can be found in groups with $N$ members - for various samples of galaxies. Fractions are plotted for groups derived as angular over-densities on the plane of the sky (i.e. using a 2D detection algorithm) in a galaxy sample drawn from the APM, the SDSS EDR and a mock catalogue (Cole et al. 1998) designed to mimic the APM survey to $b_J < 20.5$. Also plotted is the multiplicity of groups detected from the mock catalogue using a 3D technique outlined in the text, with maximal relative peculiar velocity of the galaxies set to $\Delta v = 2000\,\mathrm{km s^{-1}}$. The groups detected in two-dimensions are in good agreement for all of the galaxy samples but far fewer poor groups are detected in the mock catalogue when redshifts are also taken into account, with compensatingly more field galaxies derived. For the solid line, $f(N)$ denotes the contamination - the fraction of galaxies in a 2D group that are not in the largest 3D group that the 2d group traces. The contamintaion is calculated for $\Delta v = 2000\,\mathrm{km s^{-1}}$.

- Foremost, we are interested in knowing how many 2D groups are actually also associated in redshift-space. We define the *persistence*, $p(n)$ as the fraction of $N \geq 7$ 2D groups that trace 3D groups with $N \geq n$ members. So, if we have 100 2D groups and find 50 of them trace 3D groups with 5 or more members, $p(5) = 0.5$.

- Obviously, a single $N \geq 7$ 2D group might trace many smaller 3D groups in redshift-space. A 2D group of 16 members may actually be composed of two 3D groups of 8 members. We define the *fragmentation*, $F(n)$ as the total number of 3D groups with $N \geq n$ members that 2D groups with $N \geq 7$ trace. So, if we have 100 2D groups and find 50 of them trace 3D groups with 5 or more members but 10 of these 50 trace two 3D groups with 5 or more members, $f(5) = 0.6$. Thus, the ratio $p(n)/F(n)$ is the number of 3D groups with $n$ or more members that align along a single line of sight to produce a 2D group per 2D group. In the previous example, each 2D group with 7 or more members fragments into 1.2 3D groups with 5 or more members, or, equivalently, every fifth 2D group is composed of two 3D groups. This statistic will indicate whether a single halo is a good description of a 2D group or whether, say, several superposed sheets of matter would be a better model.

- Say we ask how many $N \geq 7$ 2D groups trace 3D groups with 5 or more members. The reverse question immediately presents itself - how many 3D groups with 5 or more members were not traced by $N \geq 7$ 2D groups? We define the incompleteness, $i(n)$, as the fraction of $N \geq n$ 3D groups that were not traced by a 2D group. We shall use the incompleteness to test whether a 3D group sample is a good reflection of $N \geq 7$ 2D groups. If the incompleteness is too high, then we are considering a 3D group sample that is not particularly representative of $N \geq 7$ 2D groups. So, if we have 100 2D groups and find 50 of them trace 3D groups with 5 or more members, so that $p(5) = 0.5$ but note that in total there are 70 3D groups with 5 or more members, then $i(5) = 20/70 = 0.29$

- Finally, as a general test of how much 3D structure our 2D groups trace, we define the *contamination*, as the fraction of galaxies in a 2D group with $N$ or more members that are not associated with the largest 3D group along that particular line of sight. So, if 2D groups with 100 or more members contain 223 galaxies in total but when considered under the criterion of Equation 4.2, only 25 of these are actually in a 3D group, then c(100) = 198/223 = 0.89. The contamination is plotted for $\Delta v = 2000\,\mathrm{km\,s^{-1}}$ in Fig. 4.1 (although it takes essentially the same value

for $\Delta v = 1000 - 4000\,\mathrm{km\,s^{-1}}$). The contamination turns out to be fairly constant with group membership. For groups with $N \geq 7$, the contamination is 0.37 (i.e. on average, 63 per cent of galaxies in a 2D group are also in the 3D group that 2D group traces). The contamination for groups with $N \geq 7$ for $\Delta v$ in the range $1000 - 4000\,\mathrm{km\,s^{-1}}$ is $0.35 - 0.39$.

Table 4.1 lists the persistence, fragmentation and incompleteness for 3D groups with $N \geq n$ members, for three different maximum allowed relative velocities between galaxies in groups ($\Delta v$ in Equation 4.2). $\Delta v = 4000\,\mathrm{km\,s^{-1}}$ roughly corresponds to the $4\sigma$ limit on relative peculiar velocities of galaxies in a ΛCDM model (Sheth & Diaferio 2001). $\Delta v = 1000\,\mathrm{km\,s^{-1}}$ roughly corresponds to the peculiar velocities observed between galaxies in densely populated regions of sky (Strauss, Ostriker & Cen 1998). We draw the following information from these tables

- The statistics are not particularly dependent on the choice of $\Delta v$, a change from $\Delta v = 1000\,\mathrm{km\,s^{-1}}$ to $\Delta v = 4000\,\mathrm{km\,s^{-1}}$ typically leads to an extra 10 per cent of $N \geq 7$ 2D groups persisting as $N \geq 6$ 3D groups. The same holds true for $N \geq 5$ and $N \geq 4$ 3D groups. We thus suggest that our simple model of having a single value of $\Delta v$ for all groups, rather than a $\Delta v$ that depends on group size and redshift, is probably accurate at about the 10 per cent level.

- It is interesting to note that the persistence is 1 for 3D groups with 2 or more members and is at the 95 per cent level for 3D groups with 3 or more members. This suggests that our 2D groups with $N \geq 7$ members are *all* equivalent to at least pairs of galaxies in redshift-space and in 95 per cent of cases, our $N \geq 7$ 2D groups are probing configurations that are at least triplets of galaxies in redshift-space. Around four-fifths of our $N \geq 7$ 2D groups have four members in 3D and two-thirds have 5 or more members in 3D.

- Comparing the persistence and fragmentation indicates how many large 2D groups are split up into several smaller groups in redshift-space. Only 1 in every 13 $N \geq 7$ 2D groups splits up into two 3D groups in redshift space. This effect, where several 3D groups arise from the same 2D group accounts for 11 per cent of the $n = 5$ 3D groups and 16 per cent of $n = 4$ 3D groups. As might be expected, 3D triplets and pairs often conspire to form large 2D groups. 30 per cent of 3D triplets that occur in a 2D $N > 7$ group do so in conjunction with other 3D groups with 3 or more

| | $n=6$ | | | $n=5$ | | | $n=4$ | | | $n=3$ | | | $n=2$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta v$ | $p(n)$ | $F(n)$ | $c(n)$ | $p(n)$ | $F(n)$ | $c(n)$ | $p(n)$ | $F(n)$ | $c(n)$ | $p(n)$ | $F(n)$ | $c(n)$ | $p(n)$ | $F(n)$ | $c(n)$ |
| $2000\,\mathrm{km\,s^{-1}}$ | 0.52 | 0.56 | 0.04 | 0.66 | 0.73 | 0.18 | 0.80 | 0.93 | 0.39 | 0.94 | 1.24 | 0.61 | 1.00 | 1.85 | 0.83 |
| $1000\,\mathrm{km\,s^{-1}}$ | 0.50 | 0.54 | 0.04 | 0.62 | 0.69 | 0.17 | 0.77 | 0.88 | 0.38 | 0.92 | 1.21 | 0.59 | 1.00 | 1.87 | 0.82 |
| $4000\,\mathrm{km\,s^{-1}}$ | 0.56 | 0.61 | 0.05 | 0.70 | 0.78 | 0.20 | 0.84 | 0.99 | 0.41 | 0.96 | 1.30 | 0.63 | 1.00 | 1.87 | 0.84 |

Table 4.1: This table records the *persistence*, $p(n)$, *fragmentation*, $F(n)$ and *incompleteness*, $i(n)$ of 2D groups with 7 or more members when compared to 3D groups with $n$ or more members. These statistics are defined in text. We show results for three different allowed maximum relative peculiar velocites between galaxies in 3D groups ($\Delta v$ in Equation 4.2). The groups are selected from a mock catalogue that mimics the APM Survey.

members. Nearly 90 per cent of 3D pairs of galaxies suffer this effect.

- About 20 per cent of 3D groups with 5 or more members were not traced by a 2D group with 7 or more members, rising to 40 per cent for 3D groups with 4 or more members and 70 per cent for 3D triplets. These figures suggest that 2D groups with $N \geq 7$ members are not tracing 3D groups with fewer than 4-5 members.

- Drawing this information together, the 3D group sample that most fairly represents the 2D $N > 7$ group sample is either 3D groups with $N \geq 5$ or with $N \geq 4$. To put this another way, although only about 45 per cent of associations in the plane of the sky with 7 or more members are also 3D groups (in redshift-space) with 7 or more members, about two-thirds of $N > 7$ 2D groups describe 3D groups with $N > 5$ and four-fifths of $N > 7$ 2D groups characterise $N > 4$ 3D groups. Many 3D triplets of galaxies are not traced by $N \geq 7$ 2D groups. Further, in 85-90 per cent of cases, $N \geq 7$ 2D groups only trace a single 3D group with more than 4-5 member galaxies. Meaning that a superposition of several sheets of matter along the line of sight is proabably not a good model for $N \geq 7$ 2D groups.

In Fig. 4.2, we attempt to illustrate the above points by plotting $N > 7$ 2D groups against $N > 7$ 3D groups and $N > 4$ 3D groups. It is reasonably easy to judge that about half of the $N > 7$ 2D groups persist as $N > 7$ 3D groups and that many of the $N > 7$ 2D structure is revealed in the $N > 4$ 3D groups, along with some $N > 4$ 3D groups that were missed.

If $N > 4$ or $N > 5$ 3D groups are a good description of the groups causing the anti-correlation signal found in Chapter 3 and if the APM mock catalogue is a good representation of galaxies in the Universe, then we are in a position to determine the cosmological implications of this section. In Chapter 3 we noted that $\Omega_m$ in galaxy groups could be estimated via

$$\Omega_m = \frac{\rho_0}{\rho_{\rm crit}} = \frac{8\pi {\rm G}}{3H_0^2}\rho_0 \tag{4.3}$$

where $\rho_0$ can be estimated as the product of the space density of galaxy groups and the average mass of a dark matter halo. This assumes that the space density of 3D groups determined in redshift-space is a good representation of proper groups. Croom & Shanks (1999) have estimated the space density of groups similar to our $N \geq 7$ 2D groups as $3 \pm 1 \times 10^{-4} h^3 {\rm Mpc}^{-3}$. The $N \geq 5$ 3D groups defined in this section, that were also present

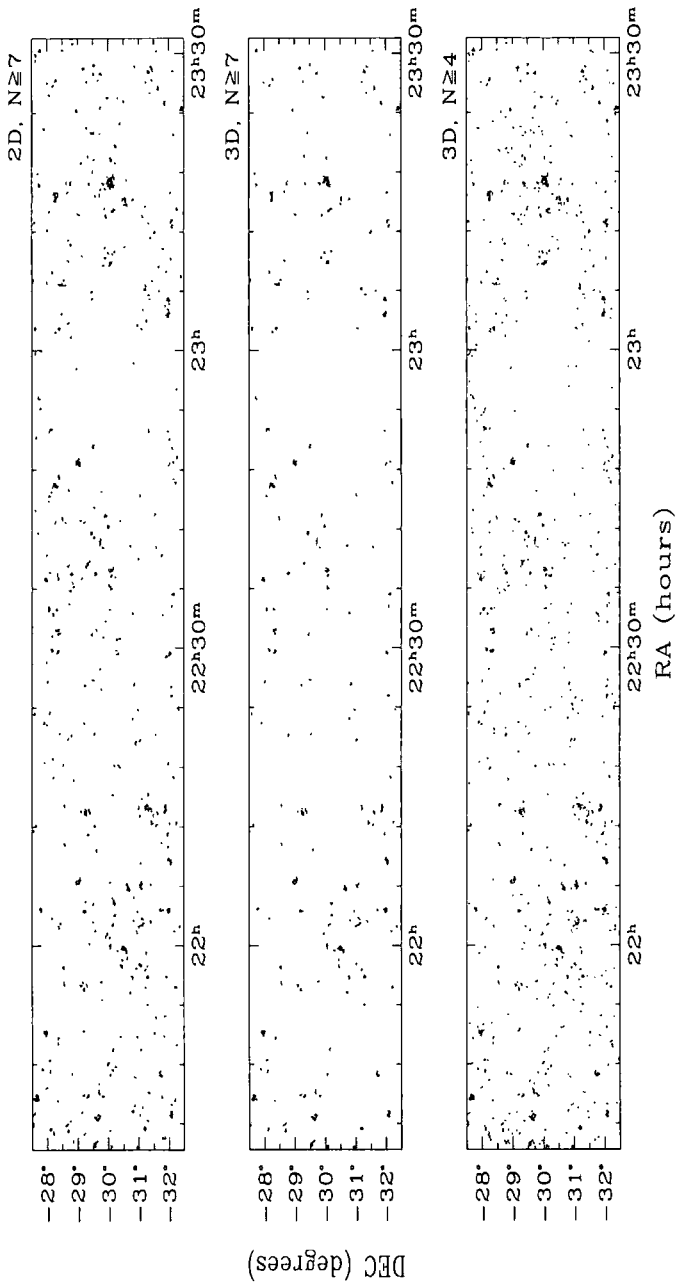


Figure 4.2: We compare structure in a mock catalogue (Cole et al. 1998) using 2D and 3D group detection. The mock catalogue is designed to mimic the APM survey to $b_J < 20.5$. 2D and 3D detection algorithms are outlined in the text. The upper panel shows groups in the mock catalogue with 7 members or more ($N \geq 7$) detected in 2D. The middle panel shows groups with $N \geq 7$ detected in 3D. Many galaxies in 2D groups are not associated in redshift-space and are thus missing from the 3D sample. The lower panel shows 3D groups with $N \geq 4$ and contains some of the missing structure from the top panel, along with some structure missed entirely by the $N \geq 7$ 2D groups.

in $N \geq 7$ 2D groups, have a sky density of 3.66 per square degree and a mean redshift of 0.126. Integrating the proper volume out to this redshift for a ΛCDM cosmology yields $3520h^{-3}\mathrm{Mpc}^3$ per square degree. The redshift distribution for APM galaxies to $b_J = 20.5$ (Baugh & Efstathiou 1993), which was displayed in Fig. 2.6 of Chapter 2, suggests that 40 per cent of the galaxy distribution is present out to the peak redshift of 0.124, so, if 40 per cent of the total volume the groups inhabit is probed to the mean redshift of 0.126, the space density of $N \geq 4$ 3D groups is then $4.0 \times 10^{-4}h^3\mathrm{Mpc}^{-3}$. This would be $4.8 \times 10^{-4}h^3\mathrm{Mpc}^{-3}$ for an Einstein-de Sitter cosmology (henceforth EdS). A similar analysis suggests a space density for our $N \geq 7$ 3D groups of $2.7 \times 10^{-4}h^3\mathrm{Mpc}^{-3}$ for a ΛCDM cosmology ($3.2 \times 10^{-4}h^3\mathrm{Mpc}^{-3}$ EdS). Thus, if it can be said that the results of Chapter 3 probe a set of 3D groups in redshift-space, they do so for a set of groups with a typical space density comparable to the figure estimated by Croom & Shanks (1999). There will be 3D groups that are associated in redshift space, however, that aren't present in our 3D group sample as they weren't dense enough to be defined as groups by our original 2D algorithm. However, to infer that the lensing properties of these less dense groupings is the same as groupings that are more compact along the line of sight is premature. Of course, it can also be argued that much of the structure picked out by our $N \geq 7$ 2D groups is not represented well by haloes of any sort but still traces mass in the Universe as projected over-densities along the line of sight. This may well be true, but would mean that the model used in Chapter 3, of the mass in $N \geq 7$ 2D groups being well represented by halo profiles is a fallacy. We will address many of these points in the next section.

## 4.3 The Cross-Correlation of 2dF QSOs and 2dFGRS Galaxy Groups

In the previous section, we noted that $N \geq 7$ groups defined as projected angular over-densities in the plane of the sky ("2D groups") are reasonably representative of $N \geq 4$ or $N \geq 5$ groups further defined by their association in redshift-space ("3D groups") but that not all of the over-densities are 3D groups, with about a quarter breaking up into pairs or triplets of galaxies in redshift-space. To better describe the statistical lensing of QSOs by galaxy groups would either require a better model than lensing through an average dark matter halo or a sample of groups that better reflects dark matter haloes. In this section, we study the cross-correlation of 2dF QSOs against galaxy groups selected

from the 2dFGRS to trace dark matter haloes (Eke et al. 2003).

### 4.3.1 2dFGRS Group Catalogue.

Eke et al. (2003) have used a percolation algorithm that is far more sophisticated than the one outlined in the previous section to define 3D groups in the full 2dFGRS survey (Colless et al. 2003). The 2dFGRS targeted APM galaxies brighter than $b_J \sim 19.45$, entirely covers the 2QZ NGC and SGC strips, and contains reliable redshifts for 90 per cent of targeted galaxies. Groups of galaxies defined by Eke et al. (2003) should be ideal for lensing analyses, as great care has been taken to designate associations of galaxies in redshift-space that best trace dark matter haloes. Eke et al. (2003) define dark matter haloes in the ΛCDM GIF N-body simulation (Jenkins et al. 1998, Kauffmann et al. 1999), produce a mock galaxy catalogue based on this simulation, then optimise the linking length between galaxies so that the returned groups of galaxies best represent the underlying (NFW) dark matter haloes. If the resulting optimised mock group catalogue is restricted to those groups with 4 or more members ($N \geq 4$), then 80 per cent of the groups contain galaxies that best describe the underlying dark matter halo. The optimal linking length is then used to pick groups out of the 2dFGRS catalogue for 2dF fields that are at least 70 per cent complete and sectors that are at least 50 per cent complete (2dF fields and sectors are described in Chapter 2. Thus the 2dFGRS group catalogue is perhaps the best representation of dark matter haloes of any group catalogue to date.

The linking length used by Eke et al. (2003) to group galaxies is actually a linking volume constrained by three free parameters (see their paper for greater detail). Eke et al. (2003) estimate the mass of any of their groups via the "gapper" estimator of (Beers, Flynn, & Gebhardt 1990). An estimate of the velocity dispersion, $\sigma_v$ of the group is calculated by ordering the group galaxies in redshift-space, calculating the velocity difference between subsequent galaxies, and then constructing a weighted mean of these velocity differences. The *rms* distance, $r$, of group members from the centre of the group is determined and the mass, which we will denote $M_{Eke}$, can then be estimated as

$$M_{Eke} = A\frac{\sigma_v^2 r}{G} \tag{4.4}$$

where $G$ is the gravitational constant. The underlying mass of the parent simulated (NFW) halo, determined from the N-body simulation, may be compared to the mass estimate of groups picked out in the mock catalogue and the value of $A$ in Equation 4.4 may

be optimised so that the estimated group masses are a good reflection of the underlying halo masses (as expected from a ΛCDM simulation). Thus, we can use the 2dFGRS group sample to compare the expected mass in simulations to the lensing mass traced by 2QZ QSOs.

The analysis in this section is carried out using a preliminary version of the 2dFGRS group catalogue, compiled in February 2003. The final version of the group catalogue, made publicly available in September 2003, was altered in several ways to make the groups picked out better reflect high mass structures in ΛCDM simulations and the centres of the groups better reflect the centres of NFW haloes. The main differences between the February and September catalogues are:

- In the February 2003 catalogue, the linking volume used to group galaxies varied with redshift but was the same for galaxies at the same redshift. The September 2003 catalogue allows the linking volume to vary at the same redshift depending on the local density of the galaxy distribution. This change acts to increase the size of larger groups but decrease the size of smaller groups, changing the mass function to better reflect that expected from ΛCDM N-body simulations.

- The centres of groups in the February 2003 version of the catalogue are calculated as the arithmetic mean of the positions of galaxies within the group. The September 2003 group-centres are defined using a procedure we shall generally refer to as the *iterative* group centre definition. The arithmetic mean of a group with $N$ member galaxies is determined, then the galaxy farthest from the mean is discarded and the arithmetic mean is recalculated for the remaining $N-1$ galaxies. This procedure is repeated $N-1$ times until only two galaxies remain, at which point the centre of the group is defined as the brightest of the two galaxies. When stacking mock groups with 10 or more members, this definition yields a group-centre that better pinpoints the NFW halo centre of the corresponding dark matter.

We do not expect these two differences to have a large effect on our results using the February 2003 2dFGRS group catalogue but we shall briefly repeat our main analysis using the published version of the group catalogue to highlight any discrepancies.

### 4.3.2 Comparison of 2dFGRS Groups and 2D APM and SDSS Groups.

In Fig. 4.3, we compare the multiplicity function of the 2dFGRS groups and groups determined from the APM mock catalogue using the 3D algorithm outlined in Section 4.1,

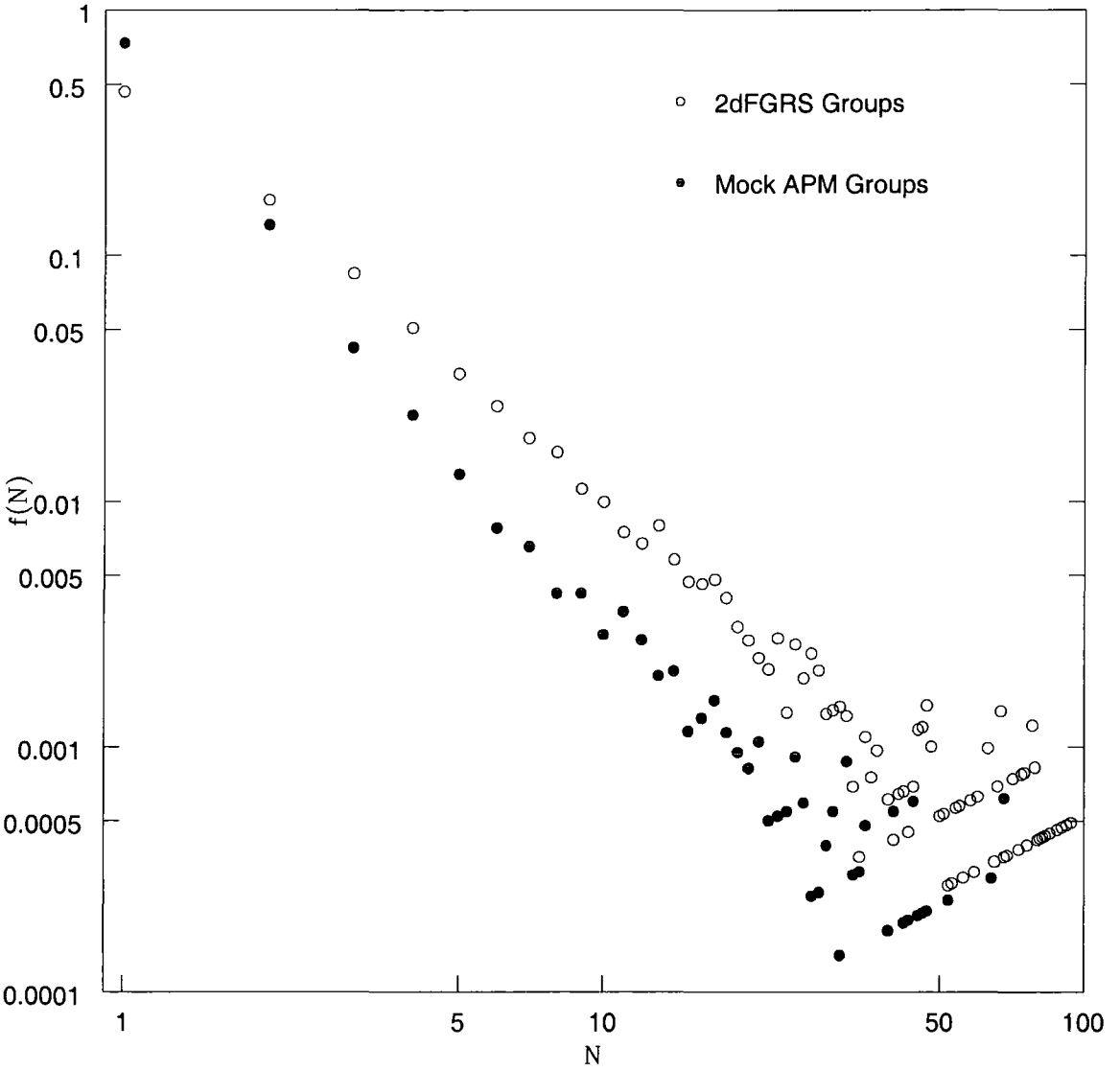


Figure 4.3: In this figure, we compare the multiplicity functions of groups from Eke et al. (2003) and groups defined in Section 4.2 from a mock APM catalogue, using a 2D algorithm coupled with the criterion that galaxies in groups have a relative peculiar velocity of no more than $\Delta v = 2000\,\mathrm{km s^{-1}}$. Although the slope of the two multiplicity functions is similar, the amplitude is very different. As Eke et al. (2003) had redshift information when formulating their group-finding algorithm, they are able to define low-redshift groups that are only marginally more dense than the background on the plane of the sky but are associated in redshift-space. This means that the group catalogue of Eke et al. (2003) generally traces lower density associations than groups defined using the Turner & Gott (1976) algorithm used in Section 4.2 and Chapter 3, above.

with $\Delta v = 2000\,\mathrm{km\,s^{-1}}$ in Equation 4.2. There are far more 2dFGRS groups and compensatingly more field galaxies assigned by our 3D algorithm. This is symptomatic of the fact that Eke et al. (2003) allowed structures that were far less over-dense (as compared to the mean background density) than allowed by our 2D group-finding algorithm to be grouped together. As the 2dFGRS groups are defined using full redshift information and at brighter magnitudes than our 3D mock APM groups, the group-finding algorithm of Eke et al. (2003) can potentially select low redshift groups that have member galaxies with relatively large angular separations. In fact, the average *rms* radius of the 2dFGRS groups is 4.7 arcminutes for groups with at least 2 members, rising to 9.0 arcminutes for groups with 7 members or more. This compares to a characteristic angular radius, determined from Equation 3.3 of Chapter 3, for our mock APM 3D groups of 0.9 arcminutes for groups with at least 2 members and 1.8 arcminutes for groups with at least 7 members.

Given that the 2dFGRS groups are on average far less compact than the mock APM 3D groups, we would not expect the 2dFGRS groups to have matches in our 2D group catalogues. Of course, our 2D groups (determined from the SDSS EDR in the 2QZ NGC strip and the APM in the 2QZ SGC strip) are also determined to $b_J \sim 20.5$, where the final 2dFGRS catalogue is only complete to $b_J \sim 18.9$ for (some) regions in the 2QZ NGC strip. However, we might expect those 2D groups that are also associated in redshift-space to be present in the 3D 2dFGRS group sample, with the proviso that the APM or SDSS galaxies included in these groups are also present in the final 2dFGRS galaxy catalogue. To test this, we determine the subsample of our APM and SDSS EDR 2D groups that have 7 or more member galaxies that *also have a matching galaxy in the final 2dFGRS galaxy catalogue.* In Fig. 4.4 and Fig. 4.5, we plot this subsample of groups against 2dFGRS groups with 7 or more members. About 45 per cent of this subsample of our groups have matches in the 2dFGRS group sample within 2 arcmin. In Table 4.2, we show the fraction of this subsample that match a 2dFGRS group with $n$ or more members, in the region of the SGC 2QZ strip, with a "match" defined as being within a separation of $1, 2$ or $5$ arcminutes. The fraction of our groups with members that have matches in the 2dFGRS groups sample is reassuringly similar to the fraction measured in the previous section when determining how many groups defined from our 2D algorithm persisted as 3D groups (see Table 4.1, above). In other words, about the expected number of our 2D APM groups turn out to be 3D groups in redshift space. Notably, as expected, about 60-80 per cent of our 2D groups with more than 4-5 members are associated in redshift-space.

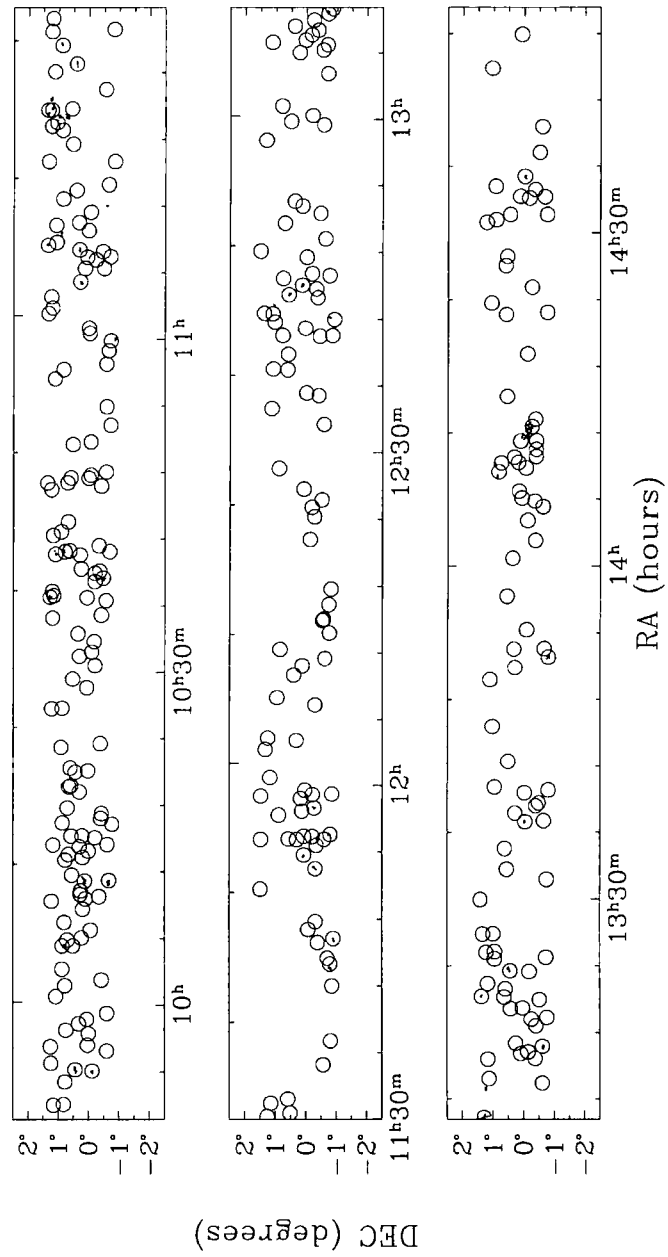


Figure 4.4: Here, we compare the groups determined in 3D by Eke et al. (2003) from 2dFGRS galaxies (plotted as circles) and groups determined as over-densities in the plane of the sky in Chapter 3 from the SDSS EDR galaxy catalogue (each galaxy in each group is plotted as a dot). The total region plotted is the 2QZ NGC strip but we only plot 2dFGRS groups that also lie within the boundaries of the SDSS EDR. We plot 2dFGRS groups that have 7 or more members. We plot SDSS EDR groups that have 7 or more member galaxies that have a match in the final 2dFGRS galaxy catalogue. About 45 per cent of SDSS EDR groups have a match within 2 arcminutes of a 2dFGRS group.

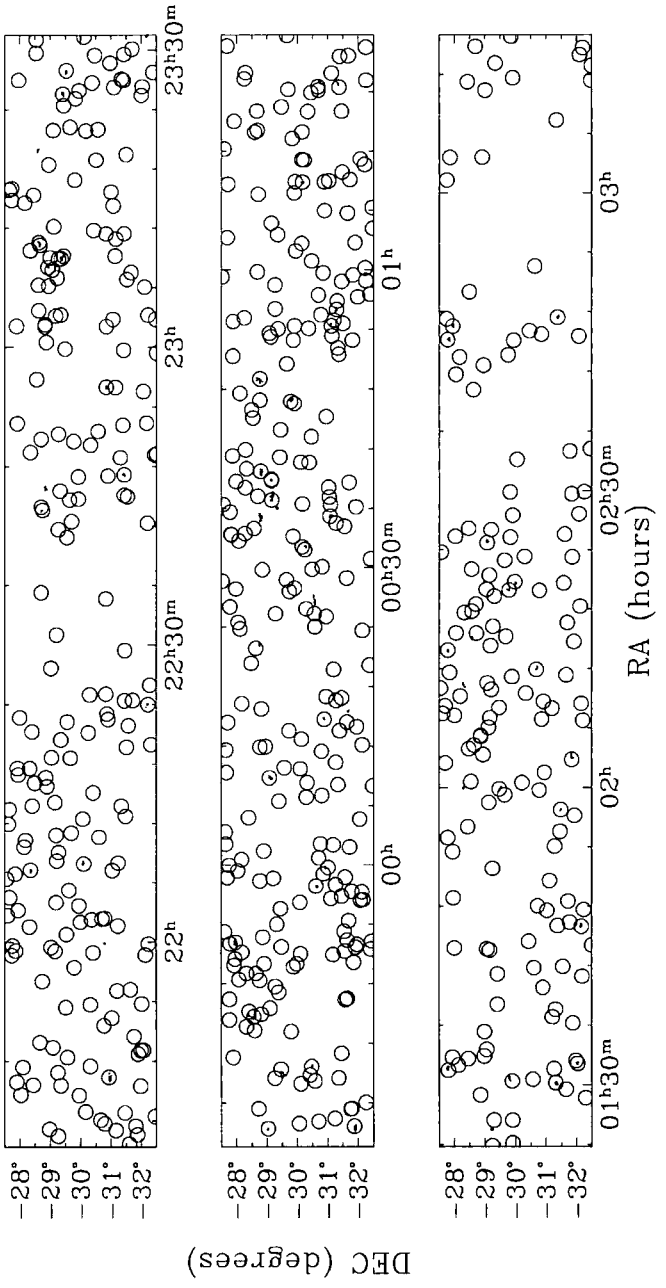


Figure 4.5: Here, we compare the groups determined in 3D by Eke et al. (2003) from 2dFGRS galaxies (plotted as circles) and groups determined as over-densities in the plane of the sky in Chapter 3 from the APM galaxy catalogue (each galaxy in each group is plotted as a dot). The total region plotted is the 2QZ SGC strip. We plot 2dFGRS groups that have 7 or more members. We plot APM groups that have 7 or more member galaxies that have a match in the final 2dFGRS galaxy catalogue. About 45 per cent of APM groups have a match within 2 arcminutes of a 2dFGRS group. Note that the void around $2^h$ $30^m$ to $3^h$ Right Ascension is not an effect of dust and is common to all of the group catalogues discussed in this thesis (see also Fig. 3.1 and discussion in Section 3.2).

| $r$ | $n = 7$ | $n = 6$ | $n = 5$ | $n = 4$ | $n = 3$ | $n = 2$ |
|---|---|---|---|---|---|---|
| $1'$ | 0.16 (0.00) | 0.16 (0.00) | 0.17 (0.00) | 0.24 (0.01) | 0.28 (0.01) | 0.33 (0.02) |
| $2'$ | 0.45 (0.01) | 0.46 (0.01) | 0.54 (0.01) | 0.61 (0.02) | 0.63 (0.04) | 0.75 (0.09) |
| $5'$ | 0.70 (0.02) | 0.72 (0.02) | 0.82 (0.03) | 0.83 (0.05) | 0.89 (0.08) | 0.97 (0.19) |

Table 4.2: This table lists the fraction of groups detected in the APM galaxy catalogue (in the region of the 2QZ SGC strip) using a 2D algorithm, that have matches (within a separation of $r$ arcminutes) in the 2dFGRS group catalogue of Eke et al. (2003). The APM groups are restricted to those that contain 7 or more galaxies ($N \geq 7$) that are found in the 2dFGRS galaxy catalogue. The fractions are shown for various sizes of 2dFGRS group ($n \geq N$ members). So, for instance, the final entry means that 97 per cent of the 2D APM $N \geq 7$ groups are within 5 arcminutes of a 2dFGRS group with at least 2 members. The number in parentheses lists the fraction of random points that would match a 2dFGRS group at this separation. This parenthetical entry is calculated as the fraction of the total area around the 2dFGRS groups considered and does not take into account the fact that clustered 2dFGRS groups will have overlapping areas (thus is actually an overestimate).

Fig. 4.4 and Fig. 4.5 also clearly show that the 2dFGRS groups contain much structure that was missed in our APM and SDSS EDR group samples. The 2dFGRS groups potentially trace large angular structures that are nevertheless associated in redshift space. In the remainder of this section we will study the cross-correlation between the 2dFGRS groups, which, we remember, are designed to trace dark matter haloes and the 2QZ QSO sample. It is unclear whether a greater lensing signal is expected in our 2D group sample, which traces lines of sight where the integrated mass is large, or the 2dFGRS sample, which traces groups of mass typical of underlying dark matter haloes in a ΛCDM cosmology, but it is clear that our models of Section 3.4 should better describe the 2dFGRS group sample than our 2D groups.

### 4.3.3 Correlation Estimator and Errors

To measure the two-point correlation function, $\omega(\theta)$, we use the estimator derived by Peebles & Hauser (1974) to study the Zwicky galaxy catalogues (Peebles 1980). This has already been discussed at length in Section 3.3, and is defined in Equation 3.5. As for Chapter 3 we produce random catalogues with 50 objects in them for every data point, created to have the same angular selection function as the considered data samples. Note that, as outlined in Section 3.3, we continue to calculate the correlation function in subdivisions that correspond to plate boundaries in the APM. This minimises the problem

of variable magnitude limits on individual plates across the APM survey, from which both the 2QZ and 2dFGRS surveys are derived.

The 2QZ QSO sample we use throughout this chapter is that of the final 2QZ catalogue (Croom et al. 2003) with a number of restrictions placed upon it to ensure we consider a fair sample of QSOs for lensing. We consider only QSOs with a redshift $z > 0.4$, to prevent the overlap in proper space of QSO and group samples. Throughout this chapter, we work with the most definitively identified QSO sample (the so-called "11" sample of Chapter 2). These restrictions in redshift and spectroscopic quality leave 21607 QSOs in the 2QZ, 12042 in the 2QZ strip in the Southern Galactic Cap (henceforth SGC) and 9565 QSOs in the strip in the North Galactic Cap (henceforth NGC). The restrictions placed upon the correlation function, namely that we only calculate the correlation function for full APM plates, lowers the number of QSOs sampled slightly, to 12039 QSOs in the SGC strip and 9237 in the NGC strip.

In Chapter 3 we introduced three forms of error that we compared the accuracy of, *Poisson Error*, *Expected Poisson Error* and *Field-to-Field Error*. We also outlined the method by which we calculate the *covariance* from Monte Carlo simulations of the data. In this chapter, we similarly create 250 Monte Carlo simulations of the 2QZ QSO sample, cross-correlate them against the 2dFGRS group sample, and compare the various error estimates to $1\sigma$ standard deviations in the values of $\omega(\theta)$ for the Monte Carlo simulations. We will additionally consider a similar estimate of error to Equation 3.8, which is essentially a weighted form of the estimate proposed by Scranton et al. (2002) when calculating the auto-correlation of galaxies in the SDSS EDR

$$\sigma_\omega^2(\theta) = \sum_{L'=1}^{N} \frac{DR_{L'}(\theta)}{DR(\theta)} (\omega_{L'}(\theta) - \omega(\theta))^2 \tag{4.5}$$

Note that $L'$ is accented in this equation. The difference between the $L'$ and $L$ in Equation 3.8 is that $L$ refers to a subsample on one of the 30 fields we subdivide the 2QZ into to estimate field-to-field error, where $L'$ refers to the subsample remaining *on the other 29 fields*. In other words, the procedure outlined in Equation 4.5 is to remove each of the fields in turn and to calculate the variance between each sample on the 29 remaining fields. We refer to this as *jackknife error*. The unweighted version of this estimate of error agrees well with simulations (see the appendix of Zehavi et al. 2002). The jackknife error estimate has an advantage over the field-to-field estimate. In Section 3.3, we noted that the field-to-field error broke down on small scales, forcing us to use an error esti-

mate based on Monte Carlo simulations of 2QZ QSOs weighted to reproduce the correct cross-correlation signal around galaxy groups. The jackknife errors *cannot* fail on small scales, provided there are data-pairs on at least two fields at these scales. The difference between the two error estimates is that when the number of data points on any single field approaches zero, the field-to-field estimate is ill-defined but the jackknife estimate remains accurate until the number of data points across the entire survey approaches zero.

In the lower panel of Fig. 4.6, we display the mean cross-correlation signal across 250 Monte Carlo realisations for the simulated 2QZ QSOs (cross-correlated against 2DFGRS galaxy groups with 7 or more members). The agreement between the NGC and SGC results are good to all scales to better than 8.5 per cent. As the expected $1\sigma$ standard error on 250 simulations is 6.3 per cent, this agreement is always better than about $1.3\sigma$. The combined Monte Carlo mean on scales from 0.5-40 arcminutes deviates from zero at most 4.4 per cent of the Monte Carlo error estimate (Only $0.70\sigma$ from expectations). However, the cross-correlation estimator for the NGC and SGC individually, clearly becomes inaccurate (for data samples of the sizes considered here) on scales less than about 1.5 arcmin. This is evidenced by the fact that the combined correlation no longer appears to be the average of the SGC and NGC results. However, the combined mean remains robust down to scales of about 0.5 arcminutes.

In the upper panel of Fig. 4.6 we plot the various error estimates outlined in this section. The error estimates appear to all agree reasonably well down to scales of about 0.5 arcminutes, the point at which we start to run out of simulated data pairs on many of the fields that the 2QZ strips are subdivided into when calculating the correlation function. In Fig. 4.7 we plot the various errors taken in ratio to the Monte Carlo error. The error estimates are all fairly good but the Poisson errors are too low across most scales. The jackknife and field-to-field error estimates remain in line with the Monte Carlo estimate across all scales. Notably, the jackknife error is always within about 7 per cent of the Monte Carlo error. Given that we expect the jackknife error to be superior to the field-to-field error on small scales, we will adopt the jackknife error as our estimator of choice in this chapter. In the lower panel of Fig. 4.7, we plot the covariance of the correlation function estimator between *adjacent* bins. The covariance is very low - less than 5 per cent on $0.5 - 10$ arcminute scales. More importantly, there is no consistent pattern in the covariance on any scale. Thus we don't expect any result we obtain to be artificially enhanced by the fact that the same data points can appear in many different data-pairs across a range of scales.

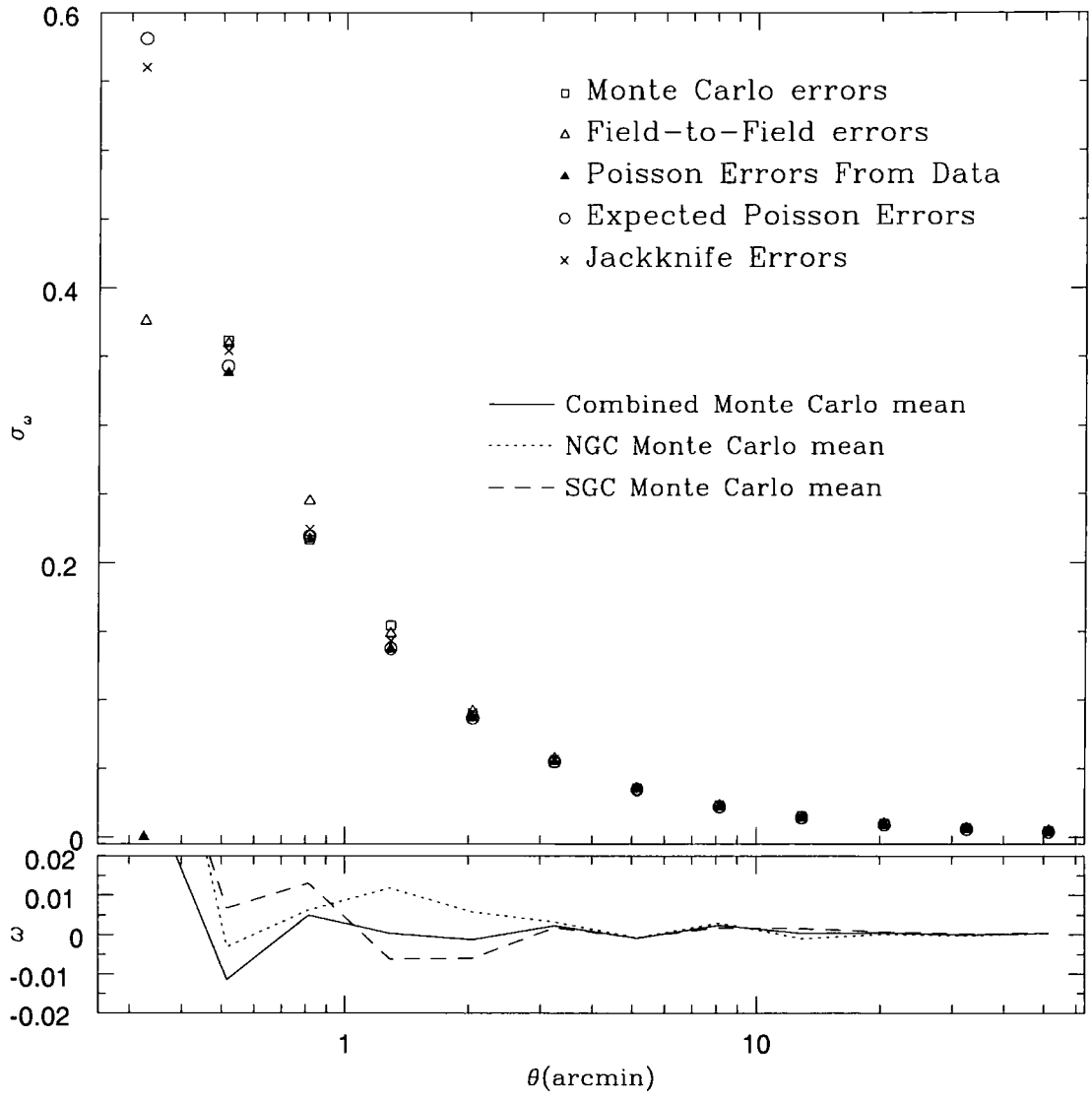


Figure 4.6: In the upper panel, we display error estimates on the correlation function for the various error estimates outlined in Section 3.3 and for the *jackknife error* defined in this section. The plotted estimates are the average errors from 250 Monte Carlo simulations of $\omega(\theta)$, calculated as the cross-correlation of Monte Carlo simulated 2QZ QSOs against 2dFGRS groups, combined for the NGC and SGC 2QZ strips. Also plotted are the Monte Carlo errors - the standard deviation in $\omega(\theta)$ of the 250 simulations. In the lower panel we plot the mean $\omega(\theta)$ of the 250 simulations for the NGC and SGC individually and for the NGC and SGC combined.

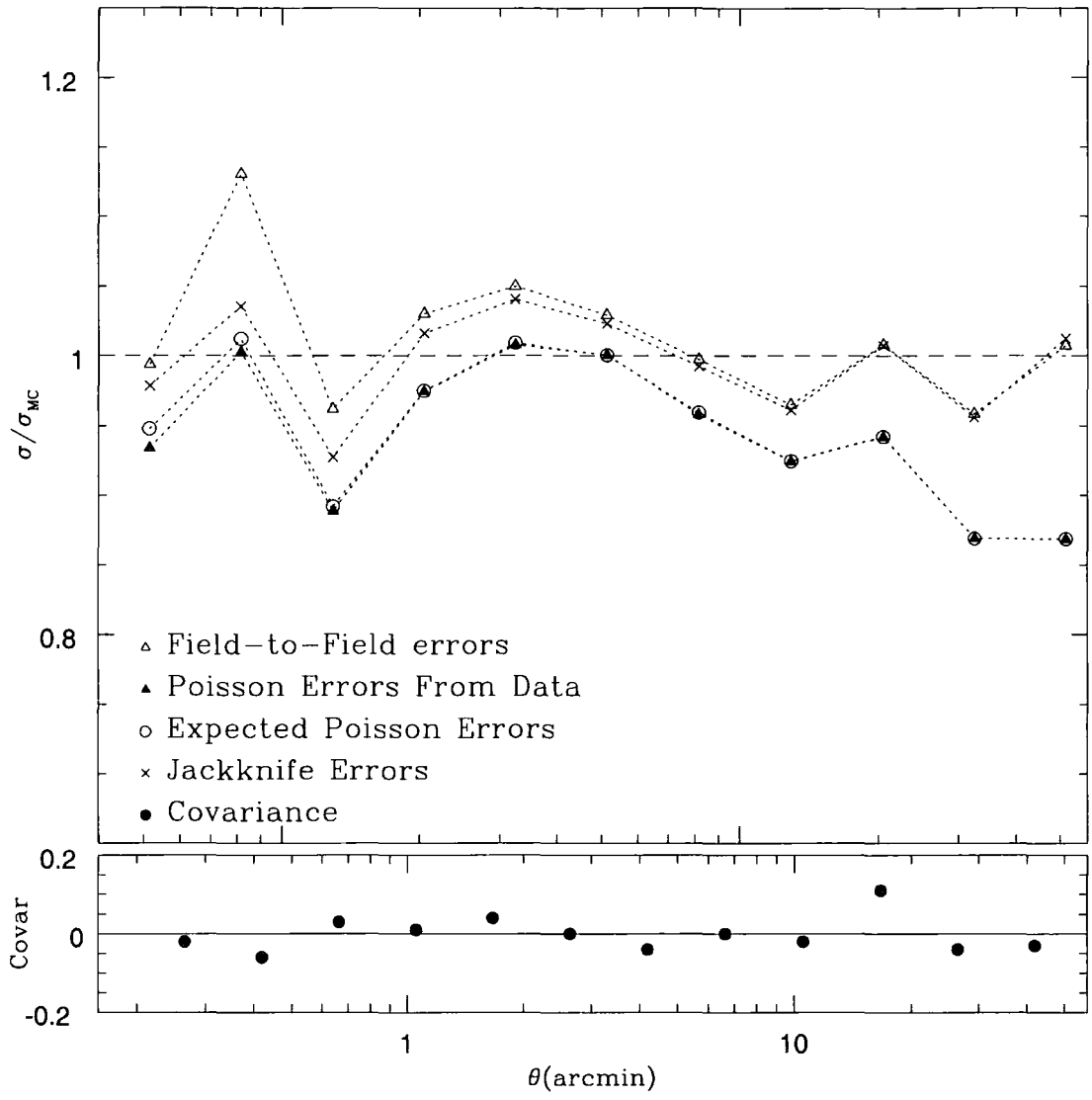


Figure 4.7: In the upper panel we display the error on $\omega(\theta)$ taken in ratio to the Monte Carlo error, for each of the error estimates outlined in Section 3.3 and for the *jackknife error* defined in this section. In all cases the errors are determined for the combined NGC and SGC sample samples. The dashed line at $\sigma/\sigma_{MC} = 1$ (where the Monte Carlo error estimate itself would lie) is drawn for comparison. The lower panel depicts the covariance between *adjacent* bins determined from 250 Monte Carlo realisations.

### 4.3.4 Cross-Correlation Analysis.

In this section, we will use the cross-correlation estimator and jackknife error tested in Section 4.3.3 to cross-correlate 2dFGRS groups and 2QZ QSOs. In Section 4.2 we determined that if the 2D groups defined in Section 3.2 trace 3D haloes, then they must trace haloes with around 5 or more members. We will use the 2dFGRS groups to test if enough lensing mass exists in 3D groups of 5 or more members to explain the strong anti-correlation found between $18.25 \leq b_J < 20.85$ 2QZ QSOs and 2D groups. We will also check whether a similarly strong, or stronger anti-correlation exists between 2dFGRS groups with any number of members and 2QZ QSOs. At the start of this section, we briefly outlined how Eke et al. (2003) optimise their groups against simulations so that the mass of the underlying NFW halo that encloses the group can be estimated. This allows us to check for trends in the cross-correlation between QSOs and galaxy groups with estimated group mass, as might be expected for a signal due to lensing. In Section 3.4 we derived a model for the cross-correlation between QSOs and galaxy groups that supposed background QSO flux was lensed by NFW haloes. We may thus compare the NFW lensing mass inferred from the group-QSO correlation signal to the mass of the simulated NFW haloes estimated for the groups by Eke et al. (2003). As much of the analysis in this section was carried out using a preliminary version of the 2dFGRS group catalogue, the last part of this section will briefly redo our analyses for the final (published) 2dFGRS group catalogue.

#### 4.3.4.1 Constraints on the Lensing Mass of 2dFGRS Groups With 5 or More Members.

In Fig. 4.8, we plot the cross-correlation between our ($z > 0.4$, "11" spectral quality) 2QZ QSO sample and 2dFGRS groups with 5 or more members ($N \geq 5$). The results are not weighted by the number of members in the group (i.e. the anaysis in this section corresponds to the "unweighted" cross-correlation functions of Section 3.5. The upper panel compares the result for the NGC and SGC strip. The lower panel shows the cross-correlation result for both possible QSO-group cross-correlation functions, centring on QSOs and comparing the distribution of galaxies to a random group distribution ($\omega_{cq}$) and centring on groups and constructing an unclustered QSO catalogue to compare with the data ($\omega_{qc}$). Note that we denote a galaxy group in the cross-correlation functions with the letter "c", for "cluster", as "g", for "group" might be construed as meaning

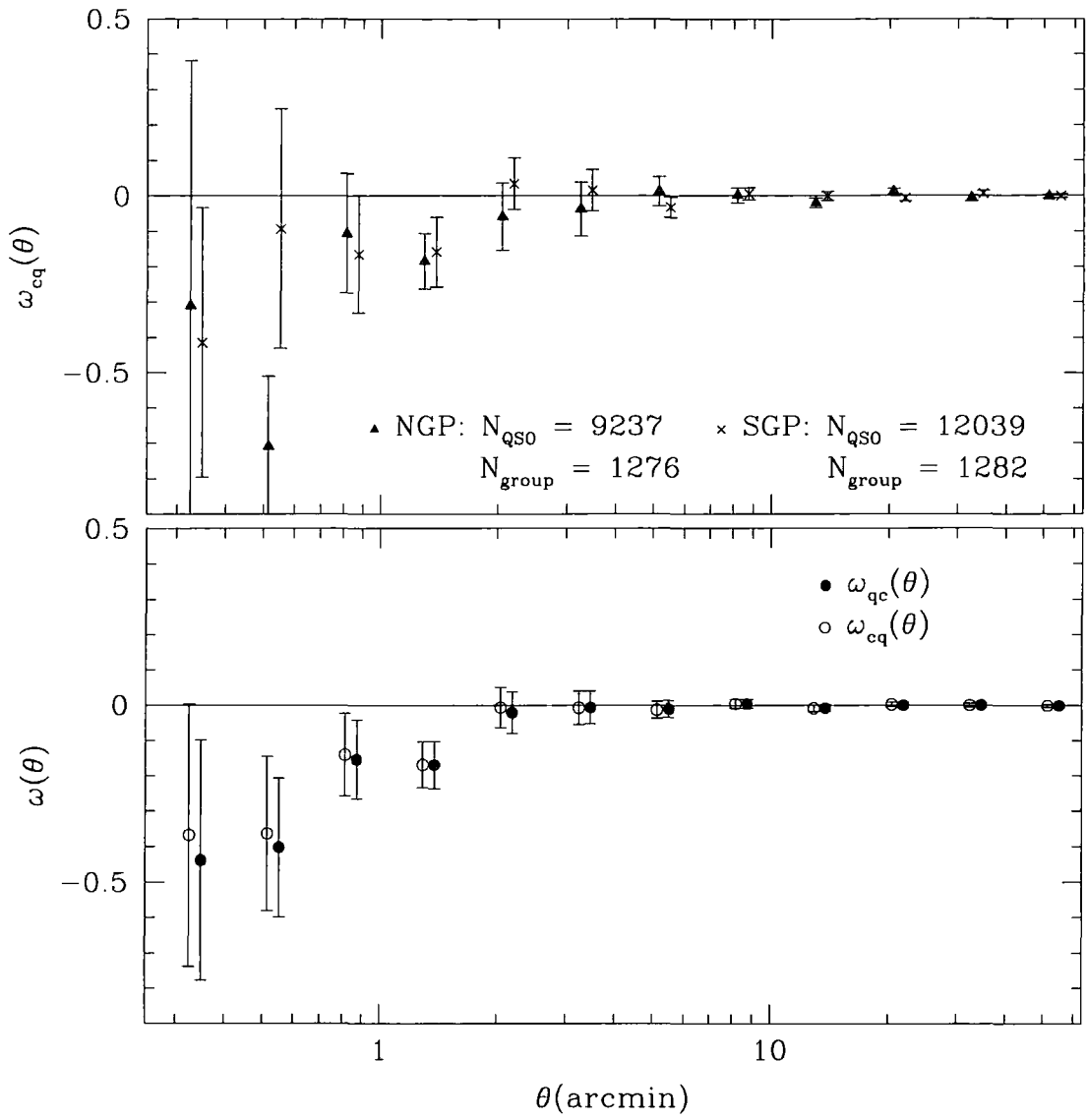


Figure 4.8: In the upper panel we display the cross-correlation between 2QZ QSOs (of "11" quality identification and redshift $z > 0.4$) and 2dFGRS galaxy groups with 5 or more members, compared for the NGC and SGC 2QZ strips. The total numbers of each object within the confines of each strip are displayed. In the lower panel, we compare the accuracy of the angular selection functions for the 2QZ and 2dFGRS by constructing the random catalogue (used to calculate the cross-correlation function, see Section 3.3) for both possible selection functions. $\omega_{cq}$ is calculated using the 2QZ masks (see Chapter 2) to select an unclustered QSO catalogue, $\omega_{qc}$ is calculated using the 2dFGRS masks (Norberg 2003) to create a random catalogue of galaxies. All errors in this plot are calculated using the jackknife estimator of Equation 4.5.

"galaxy". Our random QSO catalogue is populated using the 2QZ masks outlined in Chapter 2, to ensure the QSO distribution has the same angular selection function as the 2QZ. Our random groups are distributed according to the angular completeness mask of the 2dFGRS and populated using the 2dFGRS mask software, written by Peder Norberg and Shaun Cole (Colless et al. 2001, Norberg et al. 2002, Norberg 2003). We only populate the 2dFGRS random catalogue in fields that are at least 70 per cent complete and sectors that are at lest 50 per cent complete, as these cuts are used in defining the 2dFGRS group sample. The agreement between the cross-correlation in the NGC and SGC strips of the 2QZ is good across the range of scales considered. The agreement between the 2dFGRS and 2QZ masks is almost perfect and it is gratifying that mask software created independently for different datasets concurs so well.

In Fig. 4.9, we plot the similar result for the "11" spectral quality subsample of 2dF stars cross-correlated against ($N \geq 5$) 2dFGRS groups. Again, the agreement between the NGC and SGC results and between $\omega_{cs}$ and $\omega_{sc}$ is good. There is a discrepancy between $\omega_{sc}$ and $\omega_{cs}$, especially on large scales, similar to that noted in Section 3.5, with $\omega_{cs}$ slightly lower across most of the range. We attribute this discrepancy to the fact that stars are not really uniformly distributed across our Galaxy, having gradients (and an edge) to their distribution which is not parodied in the random catalogue. The inner points of $\omega_{sc}$ suggest that we might not trust $\omega(\theta)$ on scales less then about 0.6 arcminutes. Indeed, for both $\omega_{sc}$ and $\omega_{qc}$ the bin at 0.5 arcminues is the largest-scale bin to contain fewer than 30 data-data pairs, meaning that some of the 30 subsamples for which we calculate $\omega(\theta)$ contain no data. Note that we expected our error estimates to fail on these scales - see Section 4.3.3. Additionally, the minimum fibre placement of the 2dF instrument could effect the construction of the 2dFGRS group sample on small scales because if one group galaxy is observed, the chances of observing a second galaxy in the same group within 30 arcseconds is reduced. Using only 2dFGRS sectors that are at least 70 per cent complete in constructing the 2dFGRS group catalogue will have helped mitigate incompleteness due to minimum fibre separation effects but to be safe, we will not use bins on scales of 0.5 arcminutes or smaller when fitting models to our data, though we continue to display them in most plots.

There is a significant ($3.2\sigma$) anti-correlation between QSOs and galaxy groups on small scales, of strength $\omega_{cq}(\theta < 2 \text{ arcmin}) = -0.16 \pm 0.05$. and no significant ($1.1\sigma$) anti-correlation on the same scales in $\omega_{sc}$. We determine this significance by binning $\omega(\theta)$ in a single large bin out to the quoted scale, and determining the jackknife error in this

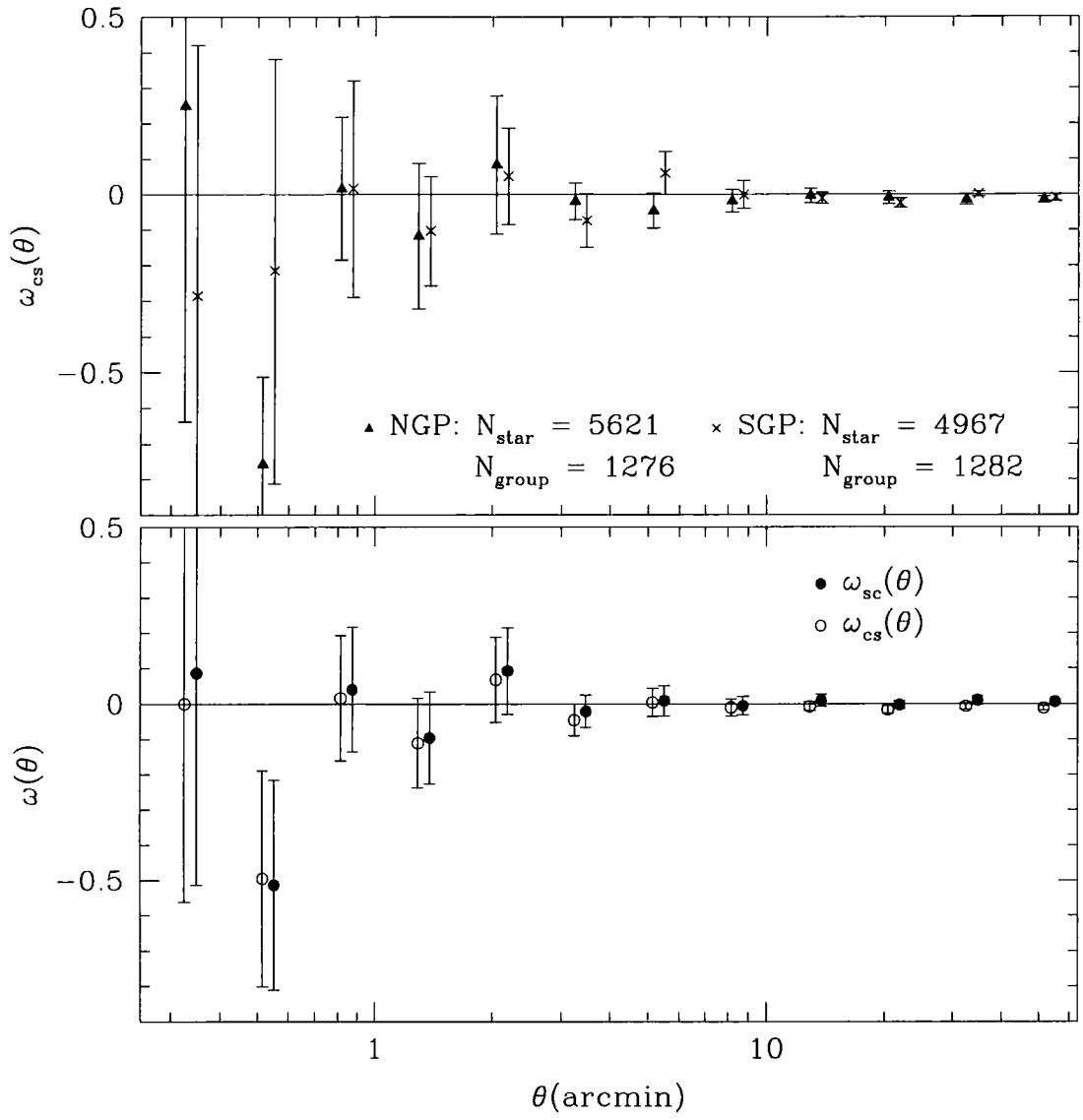


Figure 4.9: In the upper panel we display the cross-correlation between 2QZ stars (of "11" quality identification) and 2dFGRS galaxy groups with 5 or more members, compared for the NGC and SGC 2QZ strips. The total numbers of each object within the confines of each strip are displayed. In the lower panel, we compare the accuracy of the angular selection functions for the 2QZ and 2dFGRS by constructing the random catalogue (used to calculate the cross-correlation function, see Section 3.3) for both of the two angular selection functions. Errors are jackknife.

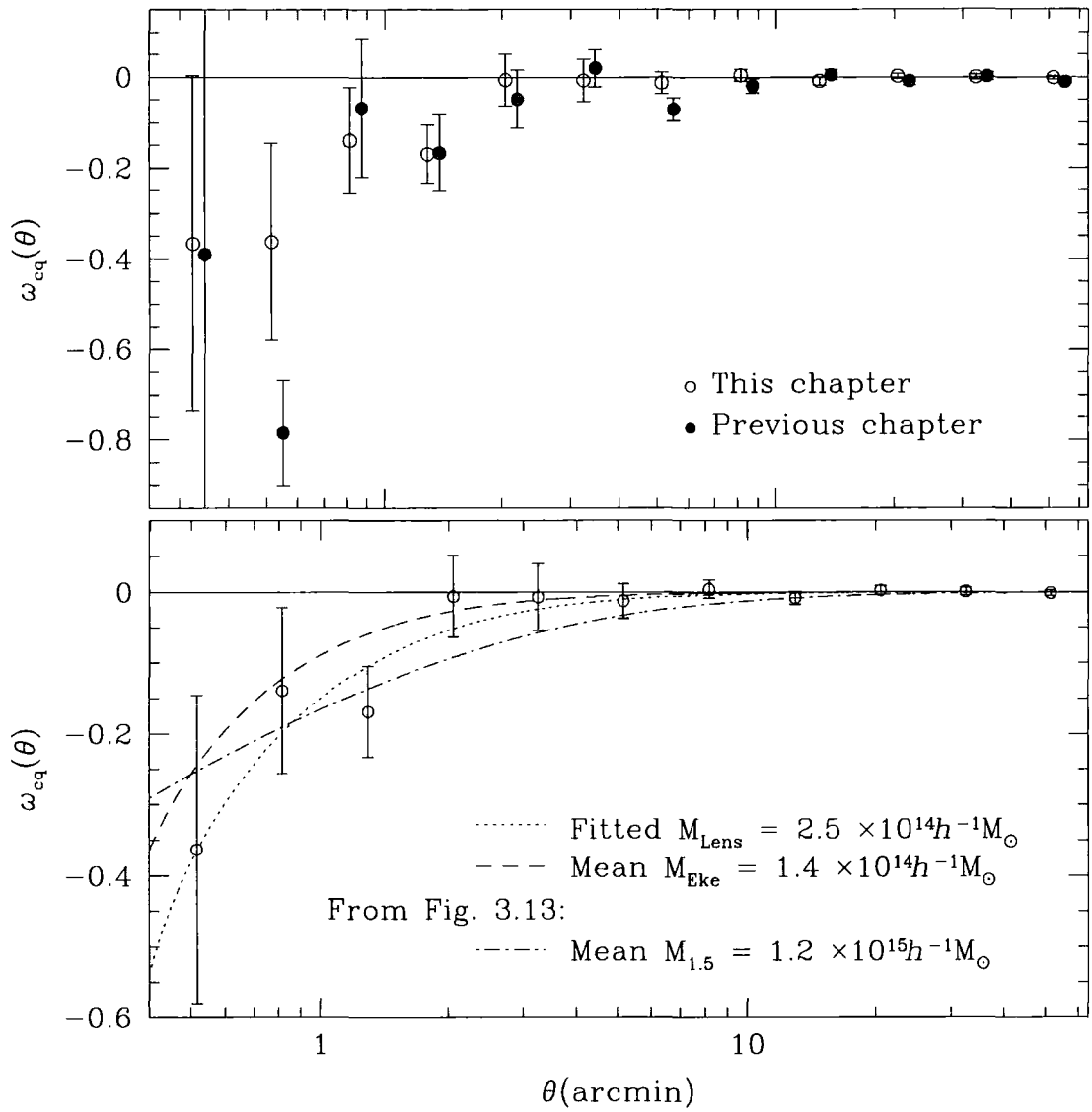


Figure 4.10: In the upper panel, we compare the anti-correlation between 2QZ QSOs and 2dFGRS galaxy groups (with 5 or more members), first plotted in Fig. 4.8, to the anti-correlation between 2QZ QSOs and groups of galaxies selected as over-densities in the plane of the sky from the APM and SDSS EDR, first plotted in Fig. 3.13. In the lower panel, we fit the NFW halo lensing model described in Section 3.4 to the anti-correlation between 2QZ QSOs and 2dFGRS galaxy groups, finding that the best-fit $\chi^2$ model has a mass of $M_{Lens} = 2.5 \pm 1.4 \times 10^{14}\, h^{-1}\mathrm{M}_{\odot}$. This compares favourably with the average mass for these groups of $M_{Eke} = 1.4 \times 10^{14}\, h^{-1}\mathrm{M}_{\odot}$, when the group masses are estimated by the method used in Eke et al. (2003). Both of these models assume the NFW radius is the mean *rms* group radius, with the *rms* radii of groups estimated as in Eke et al. (2003). The model plotted as the dot-dash line is the fitted NFW model to the 2D group data of the previous chapter. This model was first plotted in Fig. 3.15 and assumes an NFW radius of $1.5\, h^{-1}$Mpc. All errors in this plot are jackknife (see Equation 4.5)

large bin as per Equation 4.5. The significance of the anti-correlation between QSOs and groups only holds on these small scales. The 2dFGRS group-QSO anti-correlation out to 5 arcminutes is of comparable strength to the anti-correlation between 2QZ QSOs and 2D groups determined in Section 3.5, $\omega_{cq}(\theta < 5 \text{ arcmin}) = -0.034 \pm 0.033$, but is not significant on these scales and does not hold out to 10 arcminutes. In the upper panel of Fig. 4.10, we directly compare the 2dFGRS ($N \geq 5$) group-QSO anti-correlation to the (unweighted) result from Section 3.5. There seems to be excellent agreement on smaller scales but the statistical power of the two bins on 5-10 arcminute scales is sufficient to significantly reduce the overall result out to 10 arcminutes (in the sense that the 2dFGRS group result is not anti-correlated on these scales).

To highlight the difference in the result when we move from our 2D group sample of Section 3.2 to the 2dFGRS 3D group sample, we fit the NFW halo model outlined in Section 3.4 to the 2dFGRS group-QSO cross-correlation. In Section 3.4, we nominally determined the NFW mass using a scale radius of $1.5\,h^{-1}$Mpc, based on the fact that the uncovered lensing signal extended out to around 10 arcmin. Eke et al. (2003) estimate the *rms* radius of their groups as outlined at the start of this section, so in this instance we have some foreknowledge of the expected extent of the groups we probe. We will fit NFW profiles to correlation functions using the mean estimated radius of the Eke et al. (2003) groups we cross-correlate with our QSO sample. This is tantamount to replacing $1.5\,h^{-1}$Mpc in Equation 3.23 with the mean estimated group radius. Further, Eke et al. (2003) also estimate the mass of each of their groups (again see the start of this section), so we may determine the expected result for the mean estimated mass of the groups we cross-correlate against and compare the lensing mass to the expected mass from Eke et al. (2003). Note that our lensing models in this section will always use $\alpha = 0.29$ for the value of the QSO number-magnitude counts slope, as determined in Chapter 2 and used in Chapter 3. We determine the lens-source ratio $D_{ls}/D_s$ using the median redshifts of the groups ($z \sim 0.11$) and QSOs ($z \sim 1.5$), finding $D_{ls}/D_s = 0.89$ for a $\Lambda$CDM cosmology ($D_{ls}/D_s = 0.86$ for an Einstein-de Sitter cosmology - similar enough that differences in cosmology can be ignored).

In the lower panel of Fig 4.10 we fit our lensing model to the anti-correlation between $N \geq 5$ 2dFGRS groups and 2QZ QSOs. The mean *rms* radius of $N \geq 5$ 2dFGRS groups in the 2QZ region is $0.5\,h^{-1}$Mpc and their mean estimated mass is $M_{Eke} = 1.4 \times 10^{14}\,h^{-1}\mathrm{M}_{\odot}$. We use the $\chi^2$ statistic to fit the data points (on scales $0.7 < \theta < 15$ arcmin) with the anti-correlation expected for an NFW of extent $0.5\,h^{-1}$Mpc and find that the mass required

is $M_{Lens} = 2.5 \pm 1.7 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$ (reduced $\chi^2 \sim 0.5$). Interestingly, we do not rule out $M_{Eke} = M_{Lens}$, meaning there is no discrepancy between the ΛCDM simulation that Eke et al. (2003) use to calibrate their mass estimates and the mass estimate from our lensing model.

The 2dFGRS groups have a space density of $5 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, about 67 per cent higher than the space density of our 2D groups considered in Section 3.5, meaning that, although we measured $1.2 \pm 0.9 \times 10^{15} h^{-1}\mathrm{M}_{\odot}$ as the mean group lensing mass of our 2D groups in Section 3.5, the comparable estimate of the lensing mass for groups with the space density of the 2dFGRS groups is actually $7.2 \pm 5.4 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$. When we compare the measured NFW lensing mass from Fig. 4.10 to a lensing mass of $7.2 \pm 5.4 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$, we find there is no discrepancy within the allowed errors. However, a mean lensing mass as high as $7.2 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$ is ruled out by the lensing mass estimate from Fig. 4.10 at the $2.8\sigma$ level (i.e. with 99.5 per cent significance). Determining the lensing mass out to a radius of 1.5 $h^{-1}$Mpc, as in Section 3.5, instead of out to the radius implied by the 2dFGRS group radii, we determine $5.0 \pm^{4.1}_{3.4} \times 10^{14} h^{-1}\mathrm{M}_{\odot}$, which, although still not discounting $M_{Lens} = M_{Eke}$ at more than the $1\sigma$ level, would be in marginally better agreement with the results of Section 3.5. So, an important question is exactly what angular extent around groups to integrate the lensing mass out to. Fig. 4.10 would suggest that an extent of 5 arcminutes ($0.5\, h^{-1}$Mpc) is a strong upper limit on the extent of the lensing signal, in good agreement with the mean *rms* radius quoted by Eke et al. (2003) for these groups. If this *is* taken to be the extent of the lensing signal, then we cannot rule out $M_{Lens} = M_{Eke}$ at any significance but rule out lensing masses as high as the estimates given in Section 3.5 at high significance.

We have argued in Section 4.2 that if the 2D groups determined in Section 3.2 can be said to trace associations in redshift-space, then they must trace $N \geq 4$ or $N \geq 5$ 3D groups. However, the 2dFGRS group sample is a more sophisticated detection algorithm than used in Section 4.2 to construct this argument, so it might be argued the possibility remains that a different 2dFGRS 3D group membership is representative of the 2D groups. Also, if the result of Chapter 3 is a genuine probe of a particularly over-dense subset of 3D haloes, we might expect to see a strong anti-correlation between 2QZ QSOs and some subset of 2dFGRS groups. In Fig. 4.11 we show how the estimated 2dFGRS group mass, $M_{Eke}$ varies with the number of galaxies in the group (the group membership) and verify that there is an increase in mass with group membership, rather than, say groups with more members simply being low-mass but near enough that many of their members are

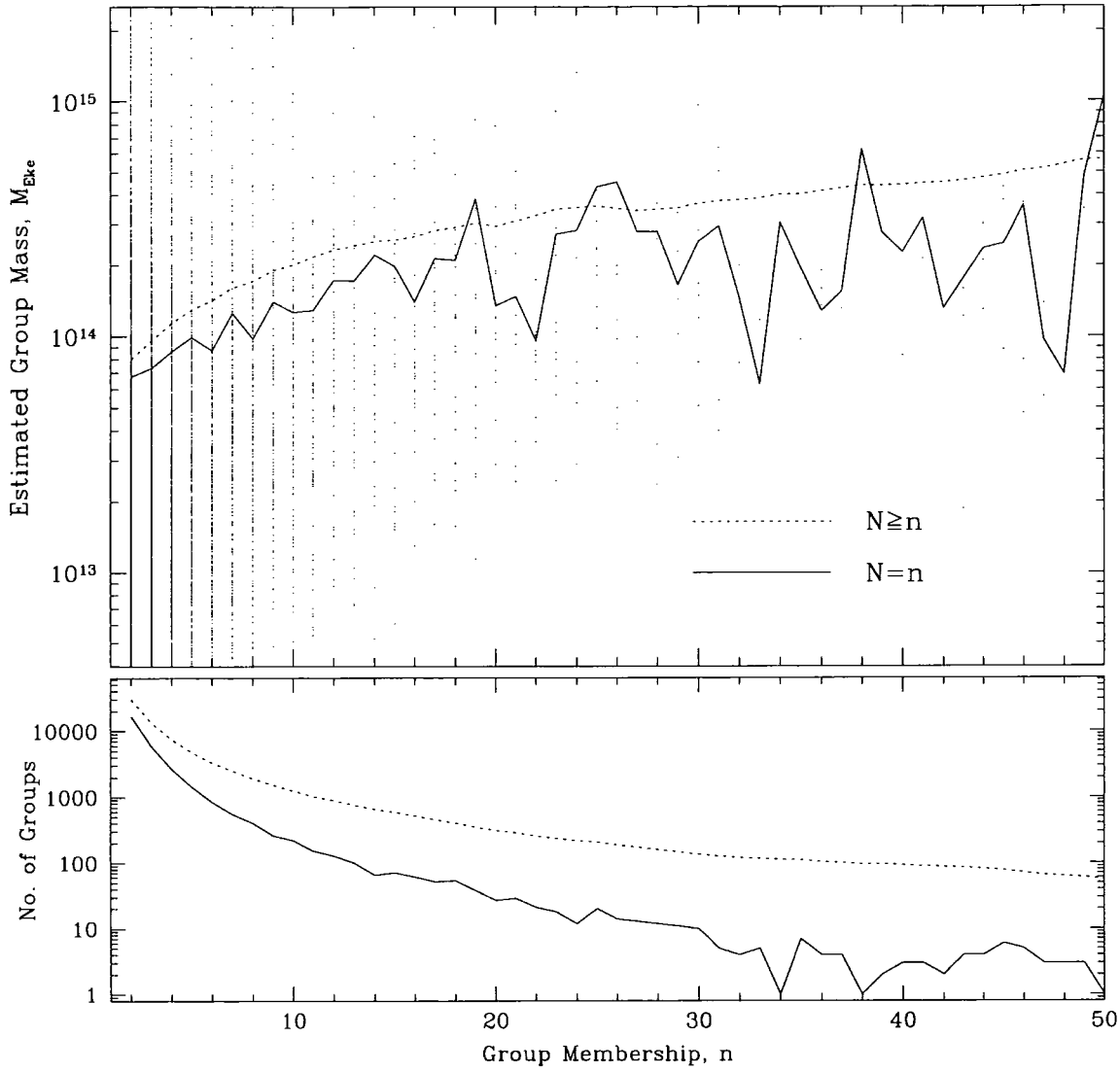


Figure 4.11: In the upper panel, we plot the mass of 2dFGRS groups as a function of the number or galaxies in a group - the group membership, $n$. The individual points are the mass and membership of each 2dFGRS group. The solid line is the mean group mass at given membership. The dashed line is the mean group mass for all groups with membership, $N$, greater than or equal to $n$. Clearly, group membership does trace group mass - the mass of groups with $N \geq 5$ is on average one-half of the mass of groups with $N \geq 15$. In the lower panel, we show the number of objects in the 2dFGRS group catalogue by group membership. The number of groups in the combined region of the 2QZ Southern strip and SDSS EDR is about 2.7 times fewer than the number plotted here.

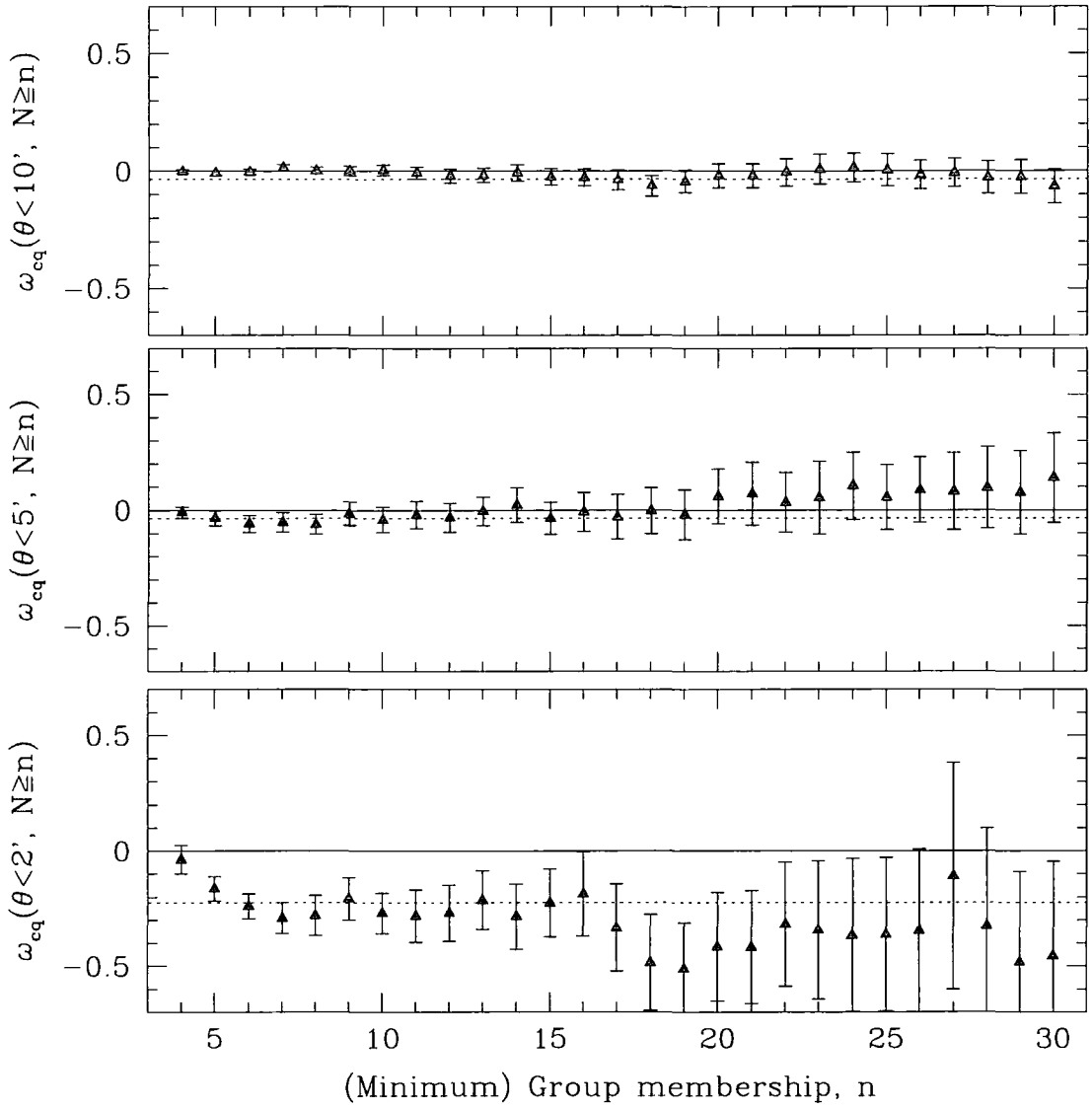


Figure 4.12: Here, we study how the cross-correlation between 2QZ QSOs and 2dFGRS Galaxy groups evolves as a function of group membership. Statistical lensing predicts a larger anti-correlation between more massive galaxy groups and faint QSOs. Each of the panels displays the cross-correlation function measured in a single bin out to a certain scale. The upper panel measures $\omega_{cq}$ in a single bin out to 10 arcminutes, the middle panel out to 5 arcminutes and the lower panel out to 2 arcminutes. The cross-correlation is measured out to this scale for galaxy groups with $N$ members, where $N \geq n$ and is plotted as a function of $n$. The dashed lines show the anti-correlation measured for the cross-correlation between 2QZ QSOs and 2D galaxy groups determined in Section 3.3. Errors in this plot are jackknife.

observed at the flux limit of the 2dFGRS.

In Fig. 4.12, we show the cross-correlation of 2QZ QSOs against galaxy groups for group membership $N \geq n$, rather than just $N \geq 5$. The maximum group membership we plot ($n = 30$) corresponds to where the number of QSO-group pairs drops below 30 in a single bin out to 5 arcminutes. Note that the jackknife error may start to underpredict the true error in $\omega(\theta < 2 \text{ arcmin})$ for $n = 15$, as data-data pairs begin to be lost on some of the subsampled 2QZ areas. A posteriori, the most significant anti-correlation in $\omega_{cq}$ is approximately where we have looked for it (n = 5) - the only more significant result is $\omega_{cq}(\theta < 2 \text{ arcmin}) = -0.241 \pm 0.054$ for n = 6, corresponding to a $4.4\sigma$ anti-correlation. Although there is generally an anti-correlation in $\omega_{cq}(\theta < 10 \text{ arcmin})$ it is never significant and is only as strong as the anti-correlation uncovered in Chapter 3 for a particular subset of group membership ($n \sim 15$). Most notably, statistical lensing predicts stronger effects for more massive groups but there is no obvious trend with group membership, except, perhaps for a marginal increase in the anti-correlation for $\omega_{cq}(\theta < 2 \text{ arcmin})$. This is a very small-scale effect, as compared to the results of Chapter 3 - we will discuss this weak trend further in Section 4.3.4.2.

We can more directly search for the expected increase in lensing signal with mass, by considering the estimated 2dFGRS masses that we have denoted $M_{Eke}$ in previous paragraphs. In Fig. 4.13, we plot the cross-correlation between our 2QZ QSO sample and two samples of 2dFGRS groups, chosen to represent the high-mass end of the 2dFGRS groups and a low-mass control sample. The high-mass subsample considers every group in the 2QZ region with $M_{Eke} > 1 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$, of which there are 657 groups with a mean estimated mass of $M_{Eke} = 3.6 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$. The low-mass subsample is chosen so that it contains about the same number of groups as the high-mass subsample. There are 641 2dFGRS groups with $M_{Eke} < 1.35 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$ and they have a mean mass of $M_{Eke} = 6.8 \times 10^{12} h^{-1}\mathrm{M}_{\odot}$. Thus, the mean difference in estimated mass between the control sample and the high-mass groups is about a factor of 50. Both the low mass and high mass groups are only considered if they have at least 5 group members (small groups have less accurate mass estimates). As previously, we fit NFW lensing models to the data to find the lensing mass, $M_{Lens}$, and compare the fit to that expected for the mean value of $M_{Eke}$. The low-mass groups have a *rms* radius of $0.3\, h^{-1}$Mpc and the high-mass groups have a *rms* radius of $0.75\, h^{-1}$Mpc. Note that the mean redshifts of the low and high-mass group samples also differ. The high-mass group sample has a mean redshift of $z \sim 0.14$, where the low-mass sample has a mean redshift of $z \sim 0.09$. However, this difference

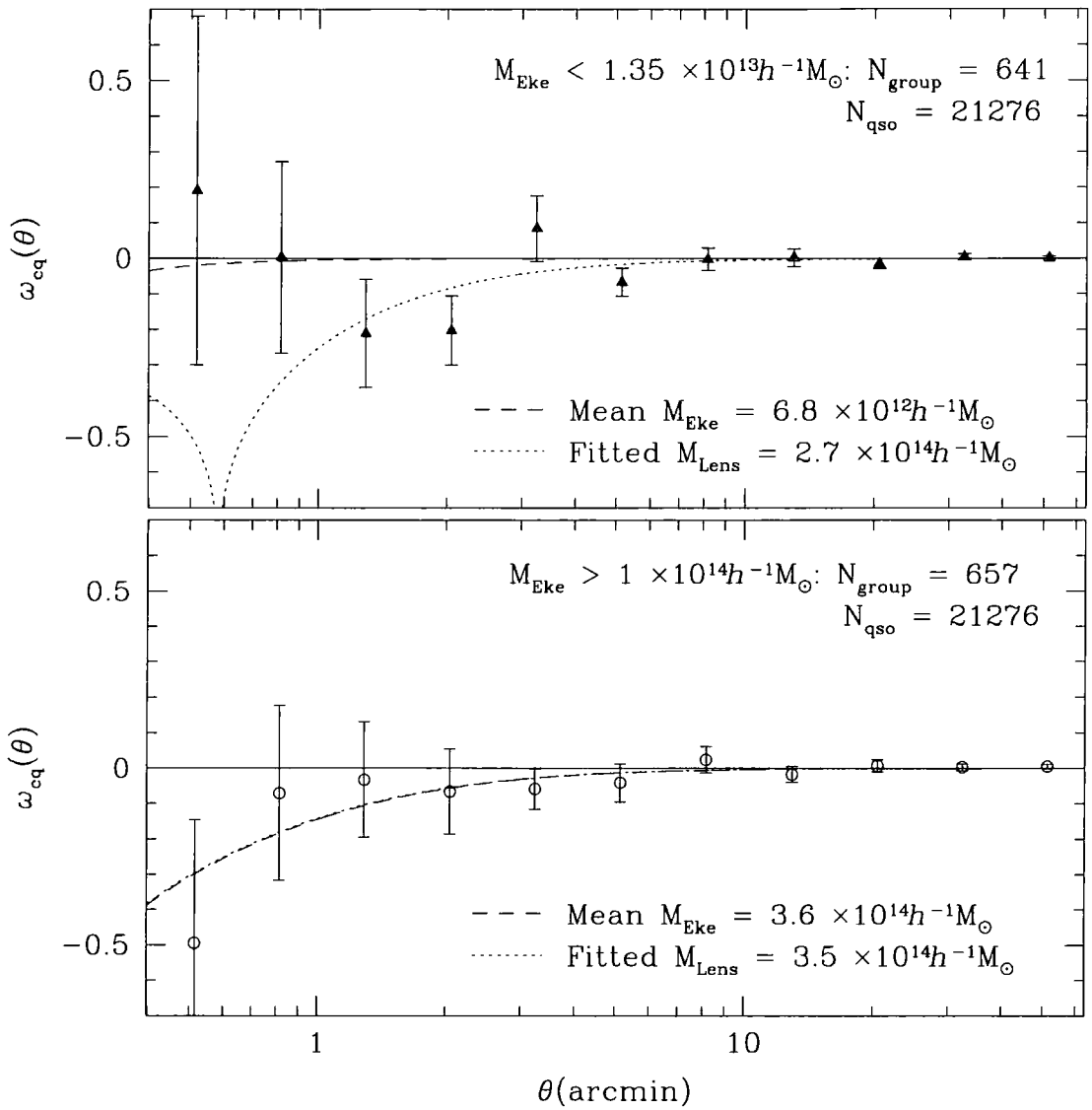


Figure 4.13: Eke et al. (2003) estimate 2dFGRS group masses as outlined at the start of this section, optimising the estimate so groups found in a mock catalogue reflect underlying halo mass. The cross-correlation between 2QZ QSOs and these groups is shown for two bins of this estimated mass, denoted $M_{Eke}$. In both panels, we fit the NFW halo lensing model derived in Section 3.4 to the data (the dotted line). We also display the NFW model that corresponds to the mean estimated mass of the haloes (the mean value of $M_{Eke}$). The best-fit lensing model in the upper panel is $M_{Lens} = 2.7 \pm^{3.1}_{2.7} \times 10^{14} h^{-1} M_\odot$ and in the lower panel is $M_{Lens} = 3.5 \pm^{4.5}_{3.5} \times 10^{14} h^{-1} M_\odot$. In neither panel can we rule out $M_{Lens} = M_{Eke}$ with any significance. The seemingly large difference in the lensing models between the panels arises as the *rms* radius of the two samples is very different (see the text). The total numbers of each object within the confines of the 2QZ are displayed. Errors are jackknife.

has almost no discernible effect on the lensing models. Above, we mentioned that we adopted values for the lens-source angular diameter distance ratio of $D_{ls}/D_s = 0.89$ (for a ΛCDM cosmology), based on a median group redshift of $z \sim 0.11$ for the lens distance. Changing this redshift to $z \sim 0.09$ or $z \sim 0.14$ results in a change of less than 2.5 per cent in $D_{ls}/D_s = 0.89$, as the source distance ($z \sim 1.5$) is so much farther away than the lens distance. We do not incorporate this small change when plotting the models in Fig. 4.13.

There is no appreciable difference in the strength of $\omega_{cq}$ between the high-mass groups and the low-mass control sample. Statistical lensing, however, predicts an appreciable difference, as evidenced by the models drawn in Fig. 4.13 for the anti-correlation expected due to lensing by a halo of mass $M_{Eke}$. In neither case do we rule out $M_{Lens} = M_{Eke}$. The lensing model fitted to the low-mass control data has difficulty finding an acceptable fit to the points - if groups this small had as much mass as predicted by the lensing fit, then the typical Einstein radius of the groups would be about 30 arcsec, which is highly unlikely given the small number of multiply lensed pairs of QSOs at such a separation (Kochanek et al. 2002, Miller et al. 2003).

#### 4.3.4.2 Constraints on the Lensing Mass of 2dFGRS Groups From the September 2003 (Published) 2dFGRS Catalogue.

Hitherto, this section has displayed results for the February 2003 version of the 2dFGRS group catalogue. The published version, circa September 2003 (Eke et al. 2003), was altered slightly by its authors; the linking volume that grouped galaxies was allowed to vary for galaxies at the same redshift to reflect the local galaxy density; and the group centre was changed from the mean of the positions of group members to an iterative definition. These changes are explained in full at the start of Section 4.3.1. In Figs. 4.14 through 4.17 we plot the results of the analysis of this section recalculated for the final version of the 2dFGRS group catalogue (Eke et al. 2003). We consider two versions of this final catalogue that have different definitions of the centre of a group. The first version uses the same definition of the group centre as the February 2003 version of the catalogue (which is also the definition we used for our group-centres in Section 4.2), namely the arithmetic mean position of the galaxies. We call this catalogue the "Arithmetic" catalogue. The second version of the catalogue uses the iterative definition of the group centre outlined in Eke et al. (2003) and in Section 4.3.1. We call this version, which corresponds exactly to the ultimately published catalogue, the "Iterative" catalogue.

The main results of the cross-correlation analysis between 2QZ QSOs and the ulti-

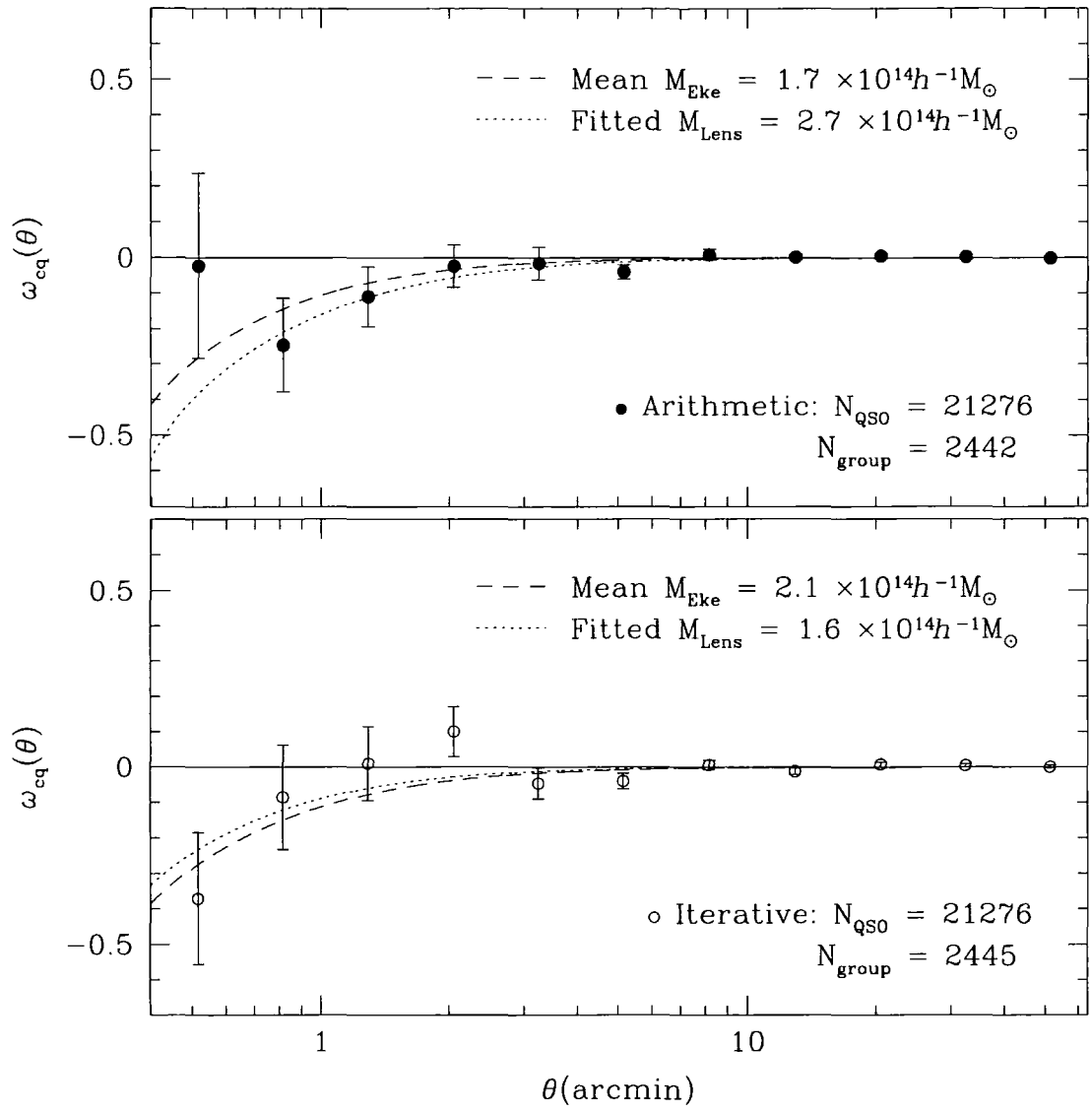


Figure 4.14: The panels in this figure are as for the lower panel of Fig. 4.10 but the analysis has been repeated using the September 2003 (published) 2dFGRS group catalogue. In the upper panel, we display the cross-correlation of 2QZ QSOs with a version of the published 2dFGRS catalogue where group-centres are defined as the arithmetic mean of the positions of galaxies in the group. The lower panel utilises a version of the catalogue where the group-centres are calculated using an iterative method outlined in the text. The best-fit values of the lensing mass are $M_{Lens} = 2.7 \pm^{1.7}_{1.6} \times 10^{14} h^{-1} M_\odot$ in the upper panel and $M_{Lens} = 1.6 \pm^{3.3}_{1.6} \times 10^{14} h^{-1} M_\odot$ in the lower panel. Errors in this plot are jackknife and the total numbers of each object within the boundaries of the 2QZ are displayed.

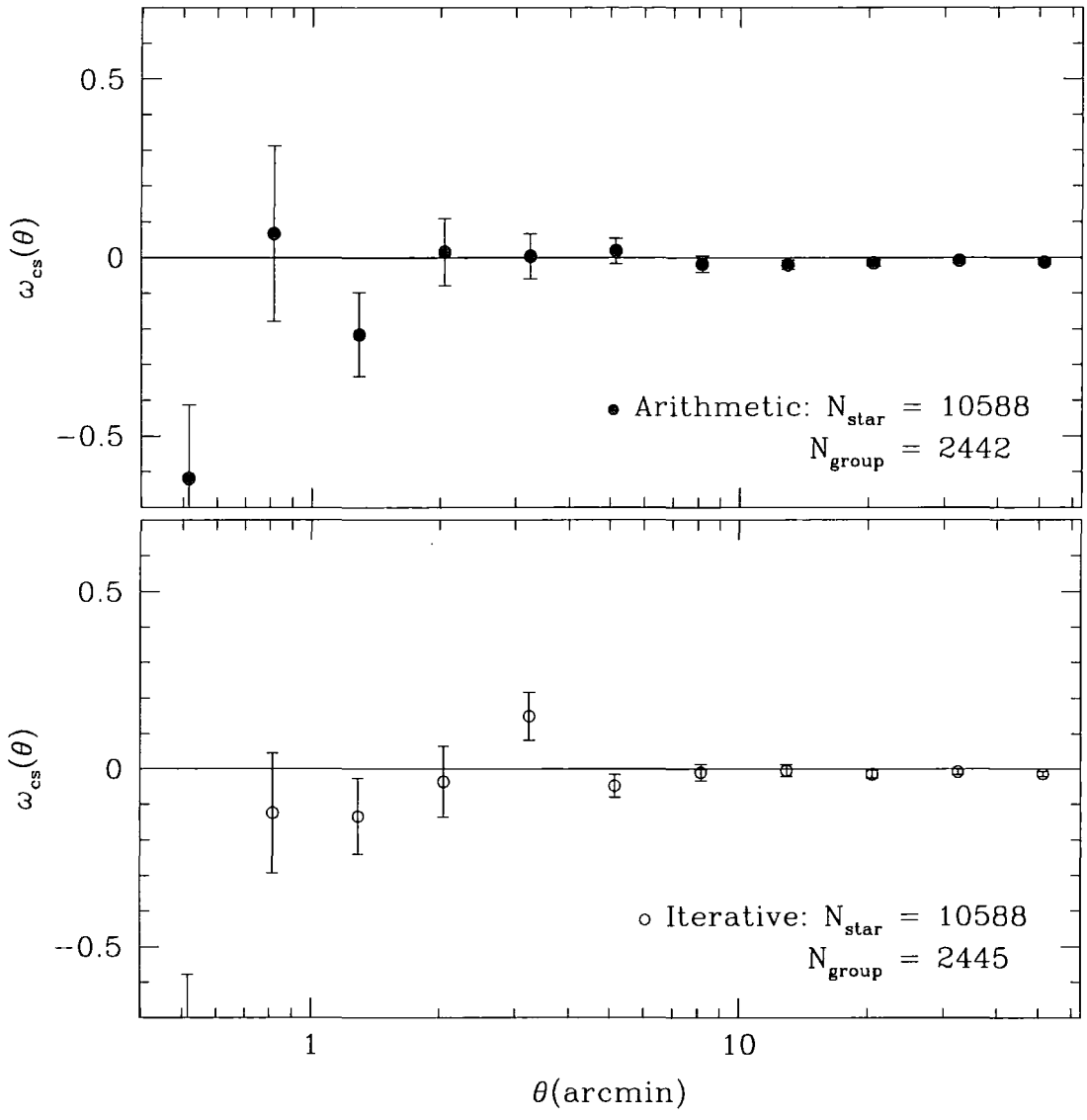


Figure 4.15: This plot displays the cross-correlation between 2QZ stars (of "11" quality identification) and 2dFGRS galaxy groups with 5 or more members, for each incarnation of the September 2003 (i.e. published) 2dFGRS group catalogue. The upper panel shows the analysis for the 2dFGRS groups with the group-centres defined as the arithmetic mean of the galaxies in the group. The lower panel shows the analysis for the 2dFGRS groups with centres defined using the iterative method outlined in the text. Errors are jackknife.

mately published version of the 2dFGRS group catalogue remain broadly the same as the for the February 2003 catalogue. However, the iterative definition of the group centre tends to remove the anti-correlation between groups and QSOs on small scales relative to the arithmetic definition, as shown in Fig. 4.8. One possible explanation for this, is that the iterative mean was defined to improve the group-centre definition for groups with 9 or more numbers - perhaps the centre of groups with 5-8 members (which are twice as large a sample as the groups with 9 or more members) is better defined by the arithmetic mean? Alternatively, as the iterative definition is designed to return groups with centres that better reflect NFW profiles in mock catalogues, it is possible that the lensing effect is hinting that the underlying mass profile is not NFW on small scales. A number of observations and theoretical considerations suggest that the NFW profile might not be a good representation of mass on small scales (Moore et al. 1999, Ghigna et al. 2000, Sand, Tommaso & Ellis 2002, Power et al. 2003).

The small-scale discrepancy between the arithmetic and iterative definition of the group centre is also apparent in Fig. 4.16, where we show the cross-correlation between QSOs and 2dFGRS groups with $n$ or more members integrated out to a range of separations.

For groups with fewer than $\sim$ 10 members, the arithmetic group-centres are significantly more anti-correlated with QSOs than the iterative definition. There is also a discrepancy between the group-centre definitions for groups with more than $\sim$ 25 members, where the iterative centres of groups are more correlated than the arithmetic centres. This is an effect of small number statistics. As in Fig. 4.12, there is a weak trend (of increasing anti-correlation with group membership) in $\omega_{cq}$ on scales of $\theta < 2$ arcmin. Any trend should be treated with caution beyond a minimum group membership of $n = 15$, as there are fewer than 30 groups of $n \geq 15$ in the data sample used. The trend of increasing anti-correlation with increasing group membership (and thus increasing mass) is consistent with a weak-lensing explanation provided the mass rises more sharply than expected from the estimated 2dFGRS group mass. Notably, if we repeat the analysis of Fig. 4.14 for groups with 15 or more members (taking arithmetic rather than iterative group-centres - although this makes little difference), we find that the fitted lensing mass is $M_{Lens} = 1.14 \pm 0.66 \times 10^{15} h^{-1}\mathrm{M}_{\odot}$, 4 times the fitted lensing mass for 2dFGRS groups with 5 or more members. The estimated 2dFGRS mass for the same groups is $M_{Eke} = 2.8 \times 10^{15} h^{-1}\mathrm{M}_{\odot}$, a value which is $1.3\sigma$ away from the lensing mass. It is possible that in larger samples of groups or QSOs this trend could significantly show that the

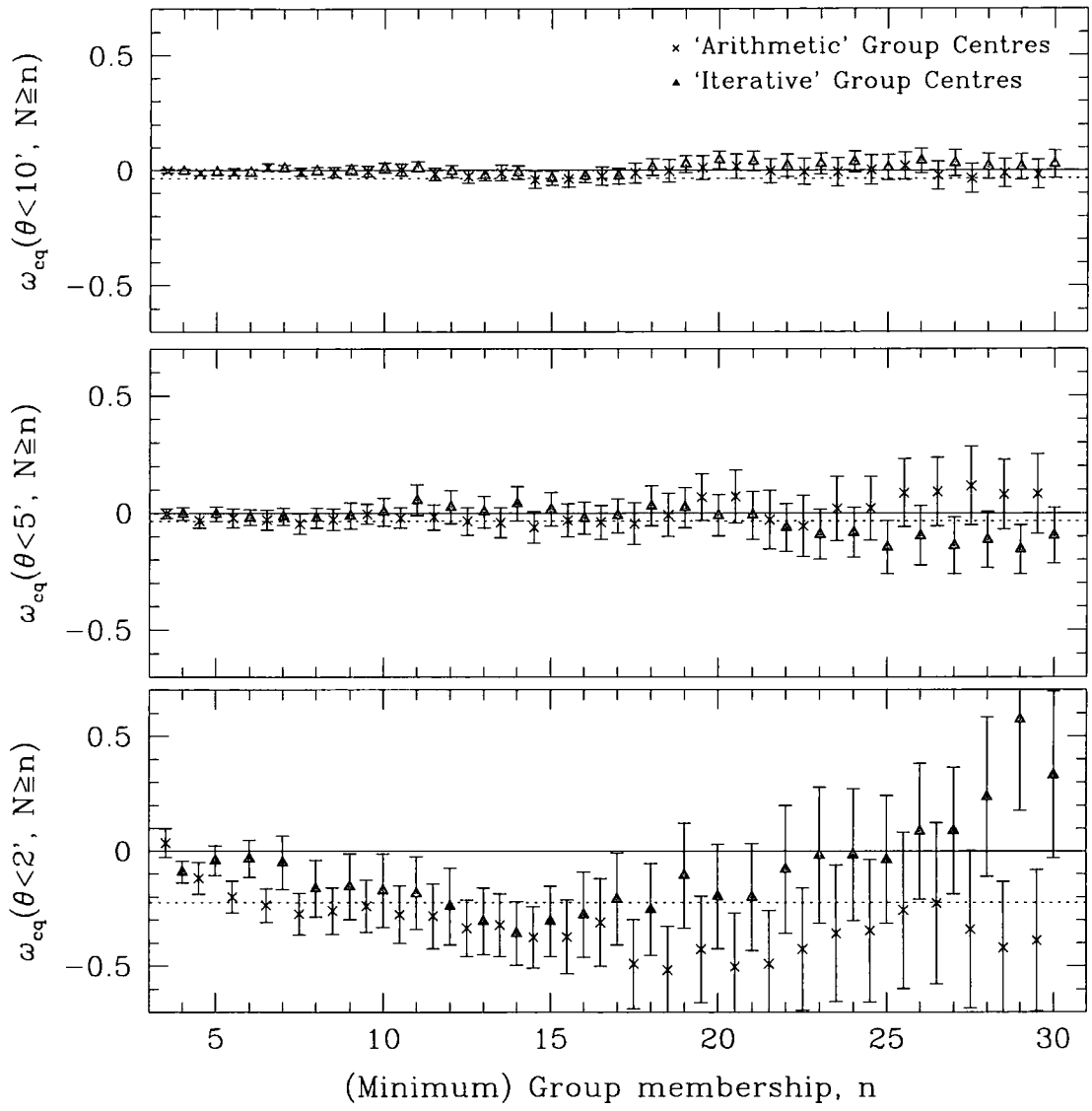


Figure 4.16: This plot is as for Fig. 4.12 but the data points are determined using the September 2003 (i.e. published) 2dFGRS group catalogue. The crosses represent the cross-correlation between groups and QSOs with the group-centres determined as the arithmetic mean of the galaxies in the group. The triangles are for the 2dFGRS groups with centres defined using the iterative method outlined in the text. The dashed lines show the anti-correlation measured for the cross-correlation between 2QZ QSOs and 2D galaxy groups determined in Section 3.3. Errors in this plot are jackknife and the crosses have been offset slightly.

lensing mass and mass from simulations are at odds. However, the current samples do not have enough statistical power to probe groups with $\sim 15-20$ or more members.

In general, whichever definition of the group centre we use, we maintain that mean lensing masses of the 2dFGRS $N \geq 5$ groups are in reasonable agreement with those estimated from simulations (i.e. we cannot rule out $M_{Eke} = M_{Lens}$ with any significance) and that the results for the September 2003 2dFGRS group catalogue are generally in agreement with those for the February 2003 catalogue. There is an increase in the lensing mass found by fitting our NFW model to the cross-correlation between the $N \geq 5$ high-mass sample ($M_{Eke} > 1.4 \times 10^{15} h^{-1}\mathrm{M}_{\odot}$) of 2dFGRS galaxy groups with "arithmetic" group-centres, as compared to the February 2003 catalogue. Comparing Fig. 4.17 to Fig. 4.13, this increase is significant at about the $1\sigma$ level - and might be expected given that the September 2003 version of the catalogue was altered to increase the number of high-mass groups picked out (as compared to the February 2003 version). This is evident in the fact that $M_{Eke}$ has increased by a factor of $\sim 30$ per cent for the high-mass sample between February and September (note also that the average *rms* group radius increased by about 15 per cent for the September 2003 catalogue as compared to the February 2003 catalogue). The fitted lensing mass of the high-mass group sample is consistent with $M_{Eke}$ in all cases.

### 4.3.5 Discussion

So, where does this leave us? We have shown in Section 4.2 that modelling the 2D over-densities defined in Section 3.4 using a single halo is largely valid provided the mass we are modelling is present in very poor groups of galaxies (with $\sim 5$ or more members). However, we have also shown that poor 2dFGRS groups determined in 3D and matched to halo profiles in N-body simulations do not show as strong an anti-correlation as 2D over-densities. For slightly richer 3D groups ($N \geq 15$) there is plausibly as strong an anti-correlation as found in the 2D groups. However, this anti-correlation is only present on very small scales ($\theta < 2$ arcmin) rather than the 10 arcminute scales found in Chapter 3 and is at the limit of the statistics that can be probed with the 2dFGRS group and 2QZ QSO samples. It is also found a posteriori - we have looked for the largest anti-correlation in our sample rather than predicitng where we should find it. It is worth discussing reasons why we might find a stronger anti-correlation between our 2D groups and QSOs than between 2dFGRS groups and QSOs, from the perspective that the results of Chapter 3 marginally favour an $\Omega_m = 1$ cosmology, where the results of this chapter

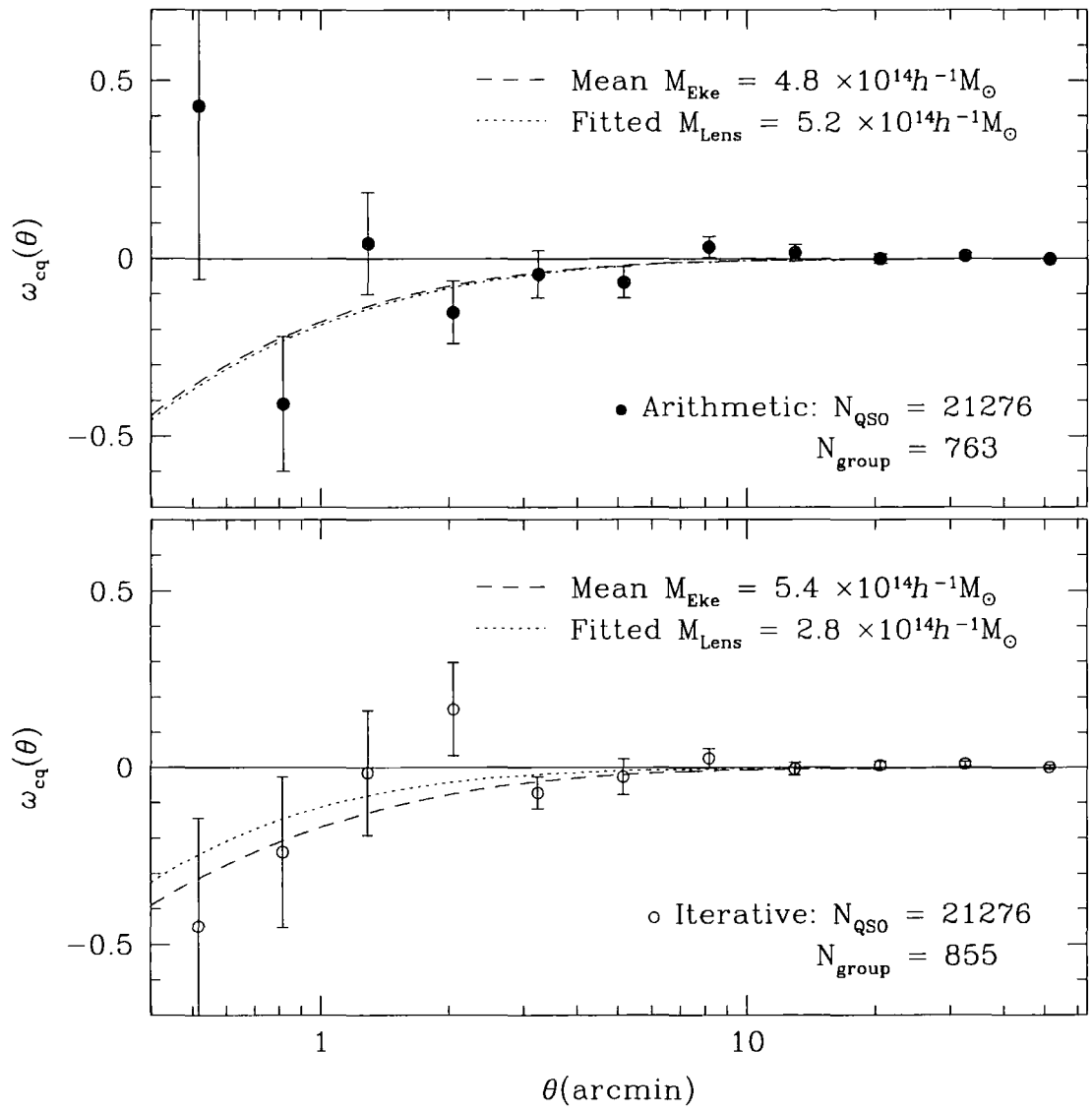


Figure 4.17: This plot is as for Fig. 4.13 but the data points are determined using the September 2003 (i.e. published) 2dFGRS group catalogue. The upper panel repeats the cross-correlation analysis when the 2dFGRS group-centres for the published groups are calculated as the arithmetic mean of member's positions. The lower panel is the analysis for published 2dFGRS groups with group-centres calculated as in Eke et al. (2003), using the iterative method outlined in the text. The fitted lensing models are $M_{Lens} = 5.2 \pm^{6.5}_{4.6} \times 10^{14}\, h^{-1} \mathrm{M}_\odot$ in the upper panel and $M_{Lens} = 2.8 \pm^{6.7}_{2.8} \times 10^{14}\, h^{-1} \mathrm{M}_\odot$. We also display the average mass calculated for the groups via the estimator from Eke et al. (2003), denoted as $M_{Eke}$. The two mass estimates agree within the (jackknife) errors. The numbers ($N$) on these plots are the totals of each object within the 2QZ boundaries.

seem to be in fairly good agreement with masses of groups drawn from simulations with an underlying ΛCDM cosmology.

Disregarding the idea that the anti-correlation between 2D associations of galaxies and 2QZ QSOs (or the lack of anti-correlation between 2QZ QSOs and 2dFGRS groups) is a statistical fluke, or hidden by the low numbers of rich groups in the 2dFGRS group sample, we might conclude that for some reason there is a stronger lensing effect caused by groups selected as 2D projections than for 2dFGRS groups. It might be argued that our 2D group sample, being representative of groups to about a magnitude fainter than the 2dFGRS 3D sample, should have more statistical power. However, as the 2dFGRS uses redshift information to pick out groups that are less over-dense on the plane of the sky than our 2D group sample, and given that our 2D group sample in the NGC, based on the SDSS EDR, only covers half of the 2QZ strip, the number of 2dFGRS groups to brighter than $b_J < 19.45$ is comparable to the number of 2D groups to $b_J < 20.5$.

It is likely that 2dFGRS groups are a fairer representation of galaxy groups than our 2D groups given the redshift information available in the 2dFGRS sample. Though the 2dFGRS groups are matched to ΛCDM simulations that, it could be argued, may or may not represent our Universe, redshift information is a useful determinant of the reality of groups despite any discrepancies between the chosen cosmology and the actual cosmology of our Universe (Diaferio et al. 1999). It is, however, possible that the groups do not trace dark matter haloes - this would require that the mock catalogue used to corroborate the 2dFGRS group-finding algorithm with an underlying N-body simulation was not a fair representation of the actual survey it mimics (which is difficult to test given that mock catalogues are usually designed to reproduce observational results determined in the survey they mimic); or that NFW haloes derived in the N-body simulation were not representative of haloes in the Universe. Despite small-scale differences between the halo profile derived both in certain simulations and observations (Moore et al. 1999, Ghigna et al. 2000, Sand, Tommaso & Ellis 2002, Power et al. 2003), and the NFW profile, we would not expect this to obviously effect the 2dFGRS group definitions. So, although the estimated mass of the 2dFGRS groups might change relative to the derived lensing masses (i.e. the constant $A$ in Equation 4.4 might change), the cross-correlation of the groups and QSOs should not be dramatically influenced. Having noted this, it seems that changing the definition of the group centre to better reflect underlying NFW profiles does have a marginal effect on the small-scale cross-correlation between groups and QSOs (as shown in Fig. 4.14), which highlights the fact that even when objective algorithms are

used to define galaxy groups, there is still a subjective element to the group sample.

As gravitational lensing traces the projected mass along the line-of-sight, it is perhaps not too surprising that 2D groups, selected as over-densities along the line-of-sight, seem to trace more mass than groups selected as isolated haloes. White, Van Waerbeke & Mackey (2002) have shown that the lensing effect of several sheets of mass along the line of sight can be significantly stronger than for isolated haloes, particularly if a smaller group lies in front of a larger cluster, in which case the mass of the foreground group may be severely overestimated. However, we have show above that even for 3D groups with as little as 5 members, few ($\sim$ 10 per cent) of our 2D groups of 7 trace more than one group along the line of sight. It might be possible that some of the smaller groups picked out in 2D are actually higher-redshift, high-mass clusters with many faint members, however, given that the 2dFGRS 3D group-finding algorithm should pick out a similar number of high mass structures as small groups, an explanation of the discrepancy in lensing mass between the two samples would require something drastic to happen to the masses of groups between the typical redshift of 2DFGRS groups ($z \sim 0.11$) and the typical redshift of our groups ($z \sim 0.13$), which seems highly unlikely.

Perhaps, then, our 2D algorithm somehow finds massive structure that the 2dFGRS 3D groups miss. Nearly all surveys of cosmological mass use structures with a greater integrated content of photon energy (many galaxies, x-ray clusters, etc.) to find structures with more mass. It is conceivable that there are dark structures in the cosmos that have a large amount of lensing mass but very little light - this would require some mechanism where a large amount of mass prevents the agglomeration of baryons (e.g. winds or shockwaves), disperses any baryons that have already formed stars (e.g. supernovae) or removes baryons entirely from view (e.g. black holes). Many of these mechanisms could be seeded by the merger of massive structures and at least some examples of massive structures with very little optical light tracing them exist (Vikhlinin et al. 1999, Miralles et al. 2002). However, it is contrived to envisage isolated, massive haloes that are better traced by projected galaxy over-densities along the line of sight than by isolated galaxy groups, meaning that modelling our 2D groups as haloes when there is no comparable signal in the 2dFGRS 3D groups they are expected to represent is almost certainly unfair. However, it is more likely that filamentary structure in the Universe (see, e.g., Plionis & Basilakos 2002) could be traced by over-densities in the projected galaxy distribution. There should not be any preferential orientation of filaments in the Universe along our line of sight but it is possible that large over-densities of galaxies along the line of sight

are a good tracer of those filaments that *are* oriented along our line of sight.

Really, then, two main points arise from this chapter's analysis. Firstly, if our 2D groups are good tracers of lensing mass, this is not because they are good tracers of 3D groups in redshift-space and not because they trace mass in dark matter haloes, hence the use of isolated haloes as a model for the mass in our 2D groups is probably not valid. Secondly, the question at hand is therefore not what our 2D groups tell us about the mass in haloes but exactly how projected luminous structure in our Universe traces underlying mass. These points will be addressed in the next chapter.

## 4.4 Conclusions

In Chapter 3 we demonstrated that a strong anti-correlation exists between faint QSOs and "2D" galaxy groups defined as projected angular over-densities on the plane of the sky and that this implies more mass in the haloes of galaxy groups than expected in an $\Omega_m = 0.3$ cosmology. In this chapter, we asked in what context the groups of galaxies defined in Chapter 3 could be considered to represent galaxy groups embedded in single haloes and went on to repeat the analysis of Chapter 3 for 2dFGRS groups of galaxies (Eke et al. 2003) designed specifically to represent underyling dark matter haloes.

We have used a mock catalogue with full redshift information (Cole et al. 1998) to show that our sample of 2D galaxy groups with 7 or more members, determined in Section 3.2, only represent 3D groups with 7 or more members in about 45 per cent of cases (though, about 63 per cent of the *galaxies* found in 2D groups are also in a 3D group). Our 2D groups are more representative of 3D groups with 4-5 or more members. We have also demonstrated that it is quite rare for our 2D groups with 7 or more members to trace several large groups along the line of sight. Even for 3D groups with as few as four members, only about 15 per cent of 3D groups align with a second group with four or more members to make a 2D configuration of 7 or more members.

We have matched our 2D groups to the February 2003, preliminary version of the 2dFGRS galaxy group catalogue (Eke et al. 2003) and shown that, when we only consider galaxies in our 2D groups that are present in the 2dFGRS, about 45 per cent of 2D groups have 3D counterparts. As these figures are in good agreement with our mock catalogue analysis, we are confident that 2dFGRS groups with 5 or more members are a reasonable 3D representation of our 2D groups, acknowledging that the 2D groups are defined from galaxy surveys that should be complete to $b_J = 20.5$ but the 2dFGRS 3D groups are only

reasonably complete to $b_J = 19.45$.

We have cross-correlated 2QZ QSOs against groups with 5 or more members from the preliminary version of the 2dFGRS catalogue. Although there is a significant anti-correlation between 2QZ QSOs and 2dFGRS groups on scales less than 2 arcminutes, that is as strong (but not as significant) as the anti-correlation between 2QZ QSOs and 2D groups, the signal is not as strong out to 10 arcminutes. This is not unique to our choice of 2dFGRS groups with 5 or more members. A posteriori, only the cross-correlation between a subset of 2dFGRS groups of $\sim 15$ or more members and 2QZ QSOs shows as significant an anti-correlation as the result for 2D groups found in Section 3.3 and this is at the limit of the statistics that can be probed by our data. It is interesting, however, that despite the lack of significance in any of the group-QSO cross-correlations, invariably faint 2QZ QSOs are anti-correlated with 2dFGRS groups, as would be expected if the groups are lensing the QSOs.

Eke et al. (2003) estimate the masses of 2dFGRS groups based on comparisons to halo masses in simulations. We have checked that the 2dFGRS group catalogue estimated mass increases, on average, with the number of galaxies in a group (the group membership) and then studied how the cross-correlation of QSOs and 2dFGRS groups changes with group membership. This test is useful, as if the 2dFGRS mass estimates are wrong, we would still expect groups with more members to be more massive and hence have a stronger lensing signal. The only trend in the group-QSO cross-correlation is an increase in the QSO-group anti-correlation with increasing group membership, from the sample of groups with 5 or more members to the sample of groups with 15 or more members. This trend is consistent with a lensing explanation, provided the lensing mass of 2dFGRS groups increases in this range more than expected from the Eke et al. (2003) 2dFGRS group mass estimates. However, again, this trend is on small scales, is not particularly significant, and is at the limits of the statistics probed by our data.

We fit the small-scale anti-correlation between 2QZ QSOs and 2dFGRS groups of 5 or more members with our NFW lensing model, derived in Section 3.4. We find no discrepancy between the fitted lensing mass and the 2dFGRS group mass estimates of Eke et al. (2003). However, taking the extent of the lensing signal to be the mean group radius estimated by Eke et al. (2003) for the 2dFGRS groups ($0.5\,h^{-1}$Mpc), which is a good upper limit on the extent of the cross-correlation between 2QZ QSOs and 2dFGRS groups of 5 or more members (5 arcminutes), we find that the large mass estimate for groups found in Chapter 3 is ruled out at high significance ($3\sigma$). However, if we fit

the lensing model using an NFW scale radius of 1.5 $h^{-1}$Mpc, as used in the analysis of Chapter 3 then the large mass estimate for groups found in Chapter 3 is marginally favoured over the Eke et al. (2003) mass estimate. Thus, the extent of the lensing signal is an important factor in discriminating between cosmological models.

Though the majority of our analysis used a preliminary version of the 2dFGRS group catalogue, the results generally hold for the ultimate, published version of the catalogue. There is some discrepancy in our cross-correlation functions when the definition of the group centre is "iterative" rather than the mean position of galaxies in the group. This discrepancy might hint at the NFW profile not being a good representation of mass on small scales, or may simply indicate the the "iterative" definition of a group centre is not particularly accurate for groups with few members.

Assuming the extent of any anti-correlation between QSOs and groups is a fair representation of the radius of the foreground group, then the only anti-correlation between 2dFGRS groups and QSOs that, a posteriori, is in agreement with the large lensing mass found for 2D groups in Chapter 3 is a subsample of 2dFGRS groups containing $\sim$ 15 or more galaxies. A comparison of the lensing mass fitted to these groups and the estimated mass from Eke et al. (2003) suggest the lensing mass is higher but only at a $1.3\sigma$ level. We conclude that the high lensing masses found in Chapter 3 for our 2D groups may not be indicative of a large lensing mass in haloes, and that the exact definition of a galaxy group has a bearing on the cross-correlation between QSOs and galaxy groups. We will proceed by attempting to implement a better model of what the relative distributions of individual galaxies and QSOs indicate about how luminous matter traces projected lensing matter.

# Chapter 5

# *The Correlation of Faint QSOs and Foreground Galaxies*

## 5.1 Introduction

In previous chapters, we have demonstrated the anti-correlation between faint QSOs and groups of galaxies and we have used gravitational lensing by simple halo models in considering the implication of this signal for cosmological parameters. Though the effect of statistical lensing should be most apparent when larger concentrations of mass, such as groups of galaxies, are considered, most recent attempts both to measure and to model associations between QSOs and foreground mass have focussed on the cross-correlation between QSOs and individual galaxies.

Seldner & Peebles (1979) were among the first authors to record a statistically significant clustering of individual galaxies with bright, high-redshift QSOs, although Tyson (1986) appears to be the first to have mentioned gravitational lensing as a possible explanation upon finding such clustering. Webster et al. (1988) developed statistical lensing as an explanation for the association of distant QSOs with foreground galaxies, suggesting that more lensing mass must be being traced than typically expected (Kovner 1989, Narayan 1989, Schneider 1989). Since then, more authors have found positive correlations between optically-selected high-redshift QSOs and galaxies (Thomas, Webster & Drinkwater 1995, Williams & Irwin 1998) and still more have found positive correlations between galaxies and distant radio-selected QSO samples (Fugmann 1988, Fugmann 1990, Bartelmann & Schneider 1994, Benítez & Martínez-González 1995, Benítez & Martínez-González 1997, Bartsch, Schneider & Bartelmann 1997, Benítez & Martínez-González 1997, Norman & Williams 2000). A large lensing effect is expected in radio-loud QSO samples where the slope of the QSO number-magnitude counts are steep.

Yet greater magnification bias may be expected in a sample that is both radio and optically selected, if the magnitudes of the QSOs in the radio and optical bands are largely independent. In such a QSO sample, *double* magnification bias may be in effect (see, e.g., Borgeest, von Linde & Refsdal 1991 and Section 5.3, below). Double magnification bias refers to the combined probability of objects that are present in (both of) two (independent) flux-limited samples being magnified into the combined sample, as compared to the unlensed probability for that sample. Very few *anti*-correlations between QSOs and individual galaxies have been detected (Benítez & Martínez-González 1997, Ferreras et al. 1997), again in samples of bright QSOs. Such anti-correlations were generally attributed to problems in optically selecting QSOs.

Over the last decade, a theoretical framework has arisen to model the cross-correlation between the density contrast of mass in the Universe and the magnification due to lensing by mass in the Universe (Bartelmann & Schneider 2001 outline this idea in length - or see Appendix C). The density contrast is simply a measure of the fractional deviation of a density perturbation from the Universal average (see Bardeen et al. 1986, or Appendix B). The nature of magnification due to lensing is discussed in depth elsewhere in this thesis (see Appendix C). The distinction between a density-magnification cross-correlation as opposed to a galaxy-QSO cross-correlation is a distinction between what we can theoretically predict and what we can practically measure. The density contrast of a distribution of lensing matter is not presently directly measurable as much of the matter is dark, but can be inferred from the arrangement of galaxies within the matter. Similarly, the magnification due to lensing may be inferred from the distribution of QSOs behind the lensing matter. This leaves us with a consistent picture where the theoretical cross-correlation of density contrast with magnification from N-body simulations of the distribution of dark matter becomes an observable cross-correlation between galaxies and QSOs (Bartelmann 1995, Dolag & Bartelmann 1997, Sanz, Martínez-González & Benítez 1997).

The contention of Zwicky (1933), that much of the matter in our Universe is non-luminous, is now well established. The primary evidence for the existence of dark matter within galactic haloes probably comes from the rotation curves of spiral galaxies, which do not drop off at large radii in check with the light but remain flat, or even rise, on the outskirts of galaxies (Rubin et al. 1985). There is strong evidence suggesting that a large fraction of the dark matter is not composed of baryons (White et al. 1993, Griffiths, Melchiorri & Silk 2001). As quantum physics informs us that luminous matter is baryonic,

we are left with a picture where dark matter may have very different properties to visible galaxies. Notably, baryonic matter interacts strongly, where non-baryonic matter must only weakly interact (or it would be more readily detectable) and this will influence the formation and distribution of dark matter relative to galaxies.

The distribution of galaxies relative to the underlying dark matter is characterised by a bias prescription (Davis et al. 1985). Kaiser (1984) demonstrated that if matter fluctuations have a Gaussian distribution, and galaxies form in regions where the fluctuations are greater than some threshold, then fluctuations in the galaxy distribution are expected to be biased tracers of dark matter fluctuations with a scale-independent amplitude (or bias parameter) $b$ (see Appendix B). This is eminently sensible on large scales, where the gravitational interaction between structures is negligible and large-scale structures are likely still imprinted with the original cosmological density field. On smaller scales matter fluctuations become non-Gaussian and baryons are subjected to various complicated mechanisms, such as supernovae-driven winds and ionisation by young, hot stars or Active Galactic Nuclei (see Adelberger et al. 2003 for an overview), which can both subdue and enhance the formation of stars and redistribute baryons relative to the underlying dark matter. As both the form of matter fluctuations and the distribution of baryons are so intricate on small scales, it is unclear exactly how the light of galaxies traces the mass in galaxies and models of non-linear galaxy bias have been developed (Fry 1986).

Whether a linear or non-linear bias model is considered, the distribution of luminous galaxies relative to the underlying mass is dependent upon bias, where the distribution of background QSOs is a direct trace of the underlying mass that lenses it. Or, in terms of what may be measured, the galaxy-galaxy auto-correlation depends doubly upon (i.e. to the square power of) the underlying mass and doubly on how strongly galaxies are biased relative to that mass but the QSO-galaxy cross-correlation, although twice a function of the underlying mass, only has a single dependence upon the galaxy bias (provided the background QSO distribution is entirely separated from the galaxy distribution). Thus, a comparison of the QSO-galaxy cross-correlation to the galaxy auto-correlation can be used to measure the strength of galaxy bias independently of the underlying mass distribution. In fact, the existence of *any* cross-correlation between galaxies and QSOs due to lensing is a confirmation that luminous matter is some biased subsample of dark matter (Bartelmann & Schneider 1993), as galaxies that were randomly distributed compared to the lensing mass could not be indicative of lensing effects.

Measurements of the strength of bias from clustering in large galaxy surveys out to

(redshift) $z \sim 0.2$ (Verde et al. 2002), comparisons of galaxy clustering with the Cosmic Microwave Background at zero redshift (Lahav et al. 2002) and weak lensing shear (Hoekstra et al. 2002) at $z \sim 0.35$ seem to be converging on a linear model of bias with the bias parameter, $b$, locally of the order of unity on scales of about $5 - 100\,\mathrm{Mpc}$ (suggesting galaxies are approximately unbiased and match the underlying dark matter distribution). However, measurements of the bias parameter from the statistical lensing of QSOs in large samples that display QSO-galaxy associations (Williams & Irwin 1998, Gaztañaga 2003) determine a much smaller bias (b $\sim$ 0.1) on small scales (suggesting galaxies trace extremely high fluctuations of dark matter on small scales). A unique measurement of galaxy bias from the cross-correlation of Gamma-Ray Bursts with QSOs (Williams & Frey 2003) suggests a similarly low value of the bias parameter. Fig 5.1 illustrates some recent determinations of the galaxy bias parameter against the scales that the measurement probed. All of these determinations of $b$ are in the redshift range $0 - 0.35$.

The 2dF QSO Redshift Survey (Boyle et al. 1999), henceforth referred to as the "2QZ" is uniquely placed to study the cross-correlation of individual galaxies with QSOs. As discussed in Chapter 3, statistical lensing predicts a positive correlation between *bright* QSOs and foreground matter and an anti-correlation between *faint* QSOs and foreground matter. The distinction between *bright* QSOs and *faint* QSOs (in this context) is made at the "knee" of the QSO number-magnitude counts, where the power-law slope of the counts is 0.4. Although significant correlations have been demonstrated between *bright* QSOs and galaxies drawn from the Sloan Digital Sky Survey Early Data Release (SDSS EDR), as yet, no statistically significant measurement of the cross-correlation of individual galaxies with a *faint* sample of spectroscopically confirmed QSOs has been made. The SDSS EDR QSO catalogue (Schneider et al. 2002) contains 3814 QSOs but only 805 are fainter than $i = 19.2$ (around the knee of the QSO number-counts). The 2QZ, on the other hand, contains 18,483 QSOs fainter than $b_J = 19.5$ (around the knee of the counts). Essentially, this means the 2QZ can probe *faint* QSO lensing statistics around 5 times more significantly than the SDSS EDR and *as* significantly as the projected (Stoughton et al. 2002) final SDSS catalogue of 100,000 QSOs.

There are additional reasons why the 2QZ is particularly appropriate for statistical lensing studies. The 2QZ contains almost as many confirmed stars as QSOs which, given that stars should not cluster relative to galaxies or QSOs, may be used as a control sample to determine if any cross-correlation effect arises from the selection of targets. The 2QZ is also ideal for demonstrating that any cross-correlation between QSOs and

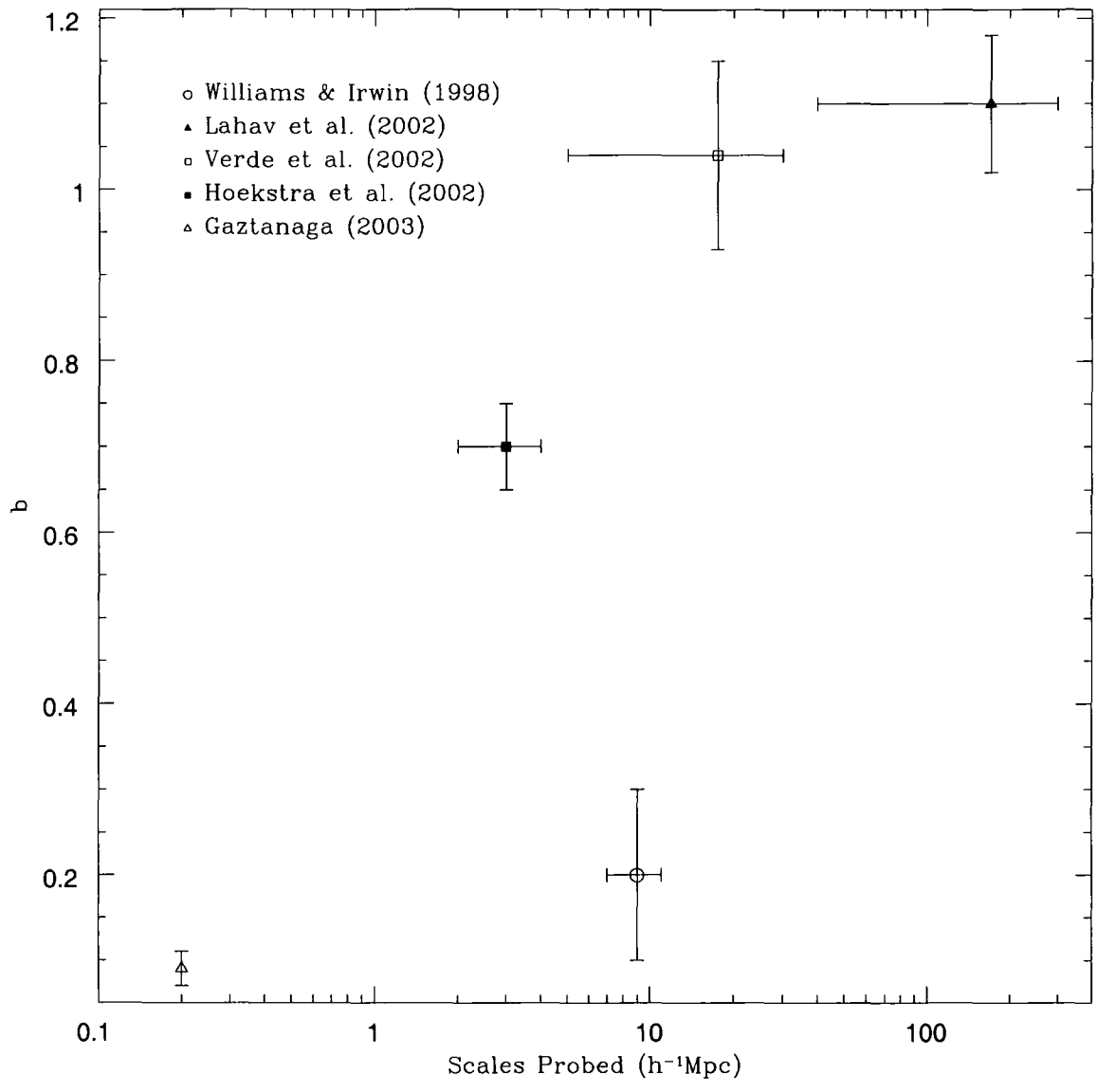


Figure 5.1: This figure illustrates recent measurements of galaxy bias by various authors using various techniques outlined in the text. The error bars display both the error on the bias measurement and the scales probed by the measurement. Where necessary for comparison, all results are shown for a ΛCDM cosmology with $H = 100h\,\mathrm{km\,s^{-1}\,Mpc} = 70\,\mathrm{km\,s^{-1}\,Mpc}$.

galaxies is most likely due to statistical gravitational lensing. The 2QZ provides QSO colours, allowing limits to be placed on how much "typical" dust in foreground galaxies could be obscuring background QSOs. The known redshifts of 2QZ QSOs allows the placing of strict limits on the QSO sample to ensure any QSOs are behind the considered sample of galaxies. Further, the 2QZ contains around 500 radio sources, allowing us to search for a cross-correlation between high-redshift, radio-loud QSOs and foreground galaxies. Although the radio sources comprise only 2 per cent of the 2QZ, they constitute a sample of spectroscopically confirmed QSOs, selected in both optical and radio passbands, comparable in size to recent samples (Bischoff & Becker 1997, Wadadekar & Kembhavi 1999). It has been suggested that double magnification bias can enhance lensing in multi-wavelength QSO samples (Borgeest, von Linde & Refsdal 1991) and observational evidence exists that this bias can be a strong statistical effect in a suitably chosen sample of even as few as 90 radio-loud QSOs (Norman & Impey 1999, Norman & Impey 2001). The radio-loud 2QZ sample can be used to investigate whether double magnification bias is expected to be a strong effect at the magnitudes typical of the 2QZ.

This chapter regards the cross-correlation between faint QSOs and foreground galaxies and its implications for cosmological parameters, particularly for galaxy bias. In Section 5.2 we outline the samples of QSOs and galaxies we shall cross-correlate and introduce the radio-loud QSO sample. Section 5.3 introduces the concept of double magnification bias and tests if it could influence our cross-correlation signal. In Section 5.4 we outline our cross-correlation methodology, use it to investigate the cross-correlation between 2QZ QSOs and galaxies drawn from the SDSS EDR and the APM Survey, and investigate possible explanations for the resulting signal. Section 5.5 introduces the models we use to investigate the cross-correlation of QSOs and galaxies in terms of statistical lensing. Section 5.6 applies lensing models to our data analysis and discusses the implications for cosmological parameters, especially how galaxies are biased relative to the underlying matter. Finally, in Section 5.7, we summarise the main results of this chapter.

## 5.2 QSO and Galaxy Samples

The QSO and galaxy samples we will cross-correlate are essentially the same as described in Chapter 4, so we will only outline them briefly here. We will introduce the radio-loud subsample of the 2QZ in slightly more length.

### 5.2.1 Entire Samples

Our QSO sample is taken from the final 2QZ catalogue (Croom et al. 2003). Unless otherwise specified, we will consider only QSOs with a redshift $z > 0.4$, to prevent the overlap in real space of QSO and galaxy samples. We shall work with only the most definitively identified QSO sample (the so-called "11" sample of Chapter 2). These restrictions in redshift and spectroscopic quality leave 21,607 QSOs in the 2QZ, 12,042 in the 2QZ strip in the Southern Galactic Cap (henceforth SGC) and 9565 QSOs in the strip in the North Galactic Cap (henceforth NGC). We shall also consider the supplemental 6dF QSO survey (henceforth 6QZ), which contains 376 QSOs after the application of our $z > 0.4$ and "11" only spectroscopic identification criteria. Completeness masks, dust correction masks and the 2dF and 6dF QSO samples have been outlined and displayed at length in Chapter 2, above.

The southern galaxy sample is taken from the APM Survey (Maddox et al. 1990b), which is considered photometrically complete to a magnitude of $b_J < 20.5$ (Maddox et al. 1990a). The northern galaxy sample is taken from the Sloan Digital Sky Survey (henceforth SDSS) Early Data Release (henceforth EDR) of June 2001 (Stoughton et al. 2002). The SDSS EDR sample is transformed into the $b_J$ band from the SDSS $g'$ and $r'$ bands using the colour equations of Yasuda et al. (2001) and cut to a magnitude of $b_J < 20.5$ to match the limit of the APM sample. Both galaxy samples are restricted to the areas in which they overlap the 2QZ strips. This leaves nearly 200,000 galaxies in the SGC 2QZ strip and nearly 100,000 in the NGC 2QZ strip. Note that the SDSS EDR only partially fills the 2QZ NGC strip.

### 5.2.2 The 2QZ Radio-Loud Catalogue

The only data used in this Section that is not described at more length in Chapter 3 is the subsample of radio-loud QSOs in the 2QZ. Radio-loud QSO candidates are selected from matches between the 2QZ UVX input catalogue (Smith et al. 1997) and discrete sources in the NRAO VLA Sky Survey (henceforth NVSS). The NVSS (Condon et al. 1998) is a survey of the entire sky north of $-40°$ Declination at 1.4 GHz frequency. Discrete sources in the NVSS were determined by fitting elliptical Gaussians to peaks of at least 2 mJy above the background. The resulting source catalogue contains 1.8 million sources and spans fluxes from around 2.5 mJy to 10 Jy. Objects from the 2QZ UVX sample that lie within 10 arcseconds of an NVSS source are considered matches (see Brotherton et al.

1998).

There are 513 radio-loud QSOs in the 2QZ with roughly equal numbers in each strip. 500 of these QSOs have a redshift z > 0.4 of which 483 are of "11" spectroscopic quality. We display all of the radio-loud QSOs in Figs. 5.2 and 5.3 with their $b_J$ magnitudes represented by circles and their 1.4 GHz flux represented by squares. The QSOs are divided into five bins in $b_J$ magnitude containing equal numbers of objects and these five bins in magnitude are plotted as different sized circles, with the brightest bin being represented by the largest circle. A similar scheme is used to display the radio flux of the QSOs, with the largest square representing the brightest bin. Note that there appears no obvious correlation between radio flux and $b_J$ magnitude. This will be investigated further in the next section, where we consider the idea that samples that are selected at more than one wavelength should be considered with special interest when discussing statistical lensing.

## 5.3 Double Magnification Bias

The idea of magnification bias (Kovner 1989, Narayan 1989, Schneider 1989), the gravitational lensing mechanism by which the distributions of galaxies and background QSOs can become correlated, has been discussed at length in this thesis (see Chapter 3). When unexpectedly large correlations between galaxies and high-redshift QSOs were first observed (Webster et al. 1988), double magnification bias (Borgeest, von Linde & Refsdal 1991) emerged as a possible explanation of the size of the correlation. Recently, Wyithe et al. (2003) have reintroduced this argument in a more general form. It is worth taking time to discuss when correlations found between foreground galaxies and QSOs are likely to be enhanced by double magnification bias. In this section, we show under simple assumptions that such an enhancement is not likely to influence 2QZ QSOs.

### 5.3.1 Background and Theory

Narayan (1989) codified how an enhancement or deficit of QSOs around foreground galaxies would be related to gravitational lensing by

$$q = \frac{1}{\mu} \frac{N(< m + 2.5 \log \mu)}{N(< m)} \tag{5.1}$$

(see Appendix C), where $q$, the enhancement factor, is the number of (lensed) QSOs expected around a galaxy compared to a random (or unlensed) QSO distribution. Here,

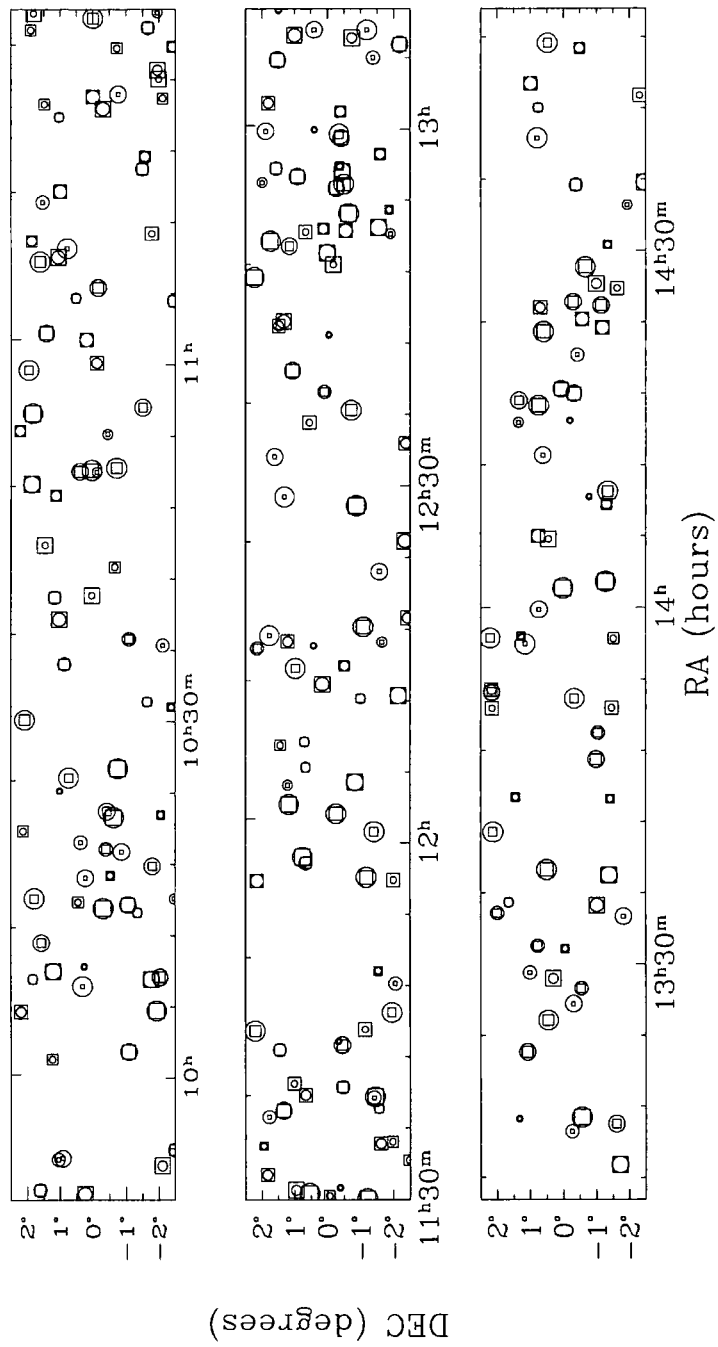


Figure 5.2: This figure displays the B1950 position of radio-loud QSOs in the 2QZ strip near the NGC. The $b_J$ magnitude of each QSO is represented by a circle, with larger circles for brighter magnitudes. The 1.4GHz radio flux of each QSO is represented by a square, with larger squares for larger fluxes. The sizes of the circles (or squares) reflects five bins in magnitude (or flux) with equal numbers of QSOs in each bin.

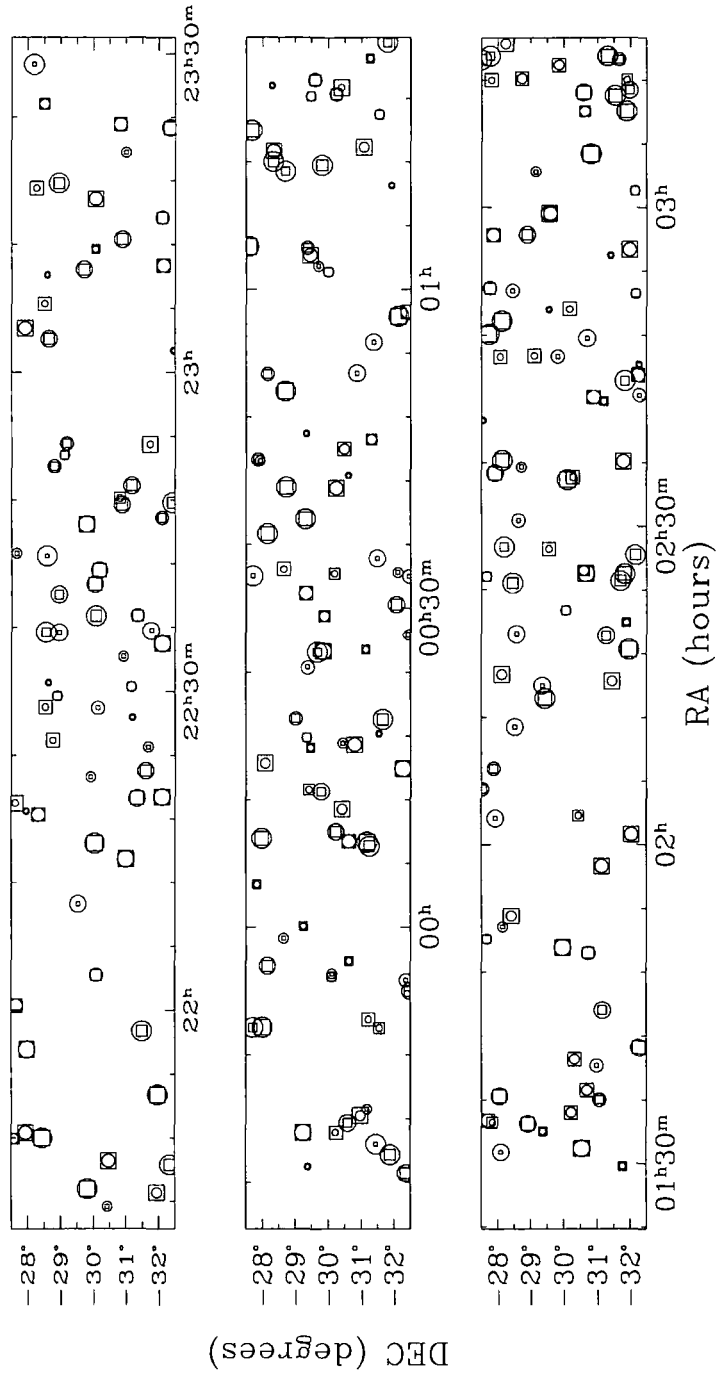


Figure 5.3: This figure displays the B1950 position of radio-loud QSOs in the 2QZ strip near the SGC. The $b_J$ magnitude of each QSO is represented by a circle, with larger circles for brighter magnitudes. The 1.4GHz radio flux of each QSO is represented by a square, with larger squares for larger fluxes. The sizes of the circles (or squares) reflects five bins in magnitude (or flux) with equal numbers of QSOs in each bin.

$\mu$ is the lensing magnification and $N(< m)$ is the number of QSOs expected brighter than magnitude $m$. The cumulative number-counts $N(< m)$ is well-approximated by a power law, with $N(< m) \propto 10^{\alpha m}$ (Boyle, Fong & Shanks 1988), so

$$q = \frac{1}{\mu} \frac{10^{\alpha(m+2.5\log\mu)}}{10^{\alpha m}} = \mu^{2.5\alpha-1} \tag{5.2}$$

Borgeest, von Linde & Refsdal (1991) noted that the same effect is true at any wavelength. This means there is a fractional enhancement of QSOs magnified into a magnitude-limited sample for each passband. The magnification should be the same in each passband, as gravitational lensing is an achromatic effect. If a QSO has magnitude $m_1$ in one passband and magnitude $m_2$ in a second passband, then, *under the assumption that the magnitude distributions in the two passbands are independent*, the enhancement factor will be

$$q = \frac{1}{\mu} \frac{N(< m_1 + 2.5\log\mu)}{N(< m_1)} \frac{N(< m_2 + 2.5\log\mu)}{N(< m_2)} \tag{5.3}$$

In general, the cumulative number-counts will not be the same in every passband but may be approximated by the power-law form $N(< m_i) \propto 10^{\alpha_i m_i}$, in each passband $i$. The enhancement factor will then become

$$q = \frac{1}{\mu} \frac{10^{\alpha_1(m_1+2.5\log\mu)}}{10^{\alpha_1 m_1}} \frac{10^{\alpha_2(m_2+2.5\log\mu)}}{10^{\alpha_2 m_2}} = \mu^{2.5(\alpha_1+\alpha_2)-1} \tag{5.4}$$

which is the same as Equation 5.2, with $\alpha = \alpha_1 + \alpha_2$.

So, the magnification bias for a subsample of QSOs that is flux limited in two (uncorrelated) passbands behaves as if it were enhancing a distribution of QSOs with a cumulative number-count power-law slope equal to the sum of the slopes in each passband. This suggests an obvious test of whether a 2QZ-foreground cross-correlation signal is due to lensing, namely, that the radio-loud 2QZ subsample might display a different signal to the parent 2QZ population. In the rest of this section we will discuss whether we might expect correlations between galaxies and the 2QZ radio-loud subsample to be enhanced by double magnification bias. We will also discuss whether the selection of 2QZ QSOs in several (highly correlated) optical bands could influence lensing enhancements.

### 5.3.2 Double Magnification Bias and the 2QZ

Condon et al. (1998) have demonstrated that the cumulative number of NVSS radio sources of given flux, $S$, is well approximated by a power law of the form $N(> S) \propto S^{-\beta}$.

This is equivalent to our usual power law form for magnitudes of $N(< m) \propto 10^{\alpha m}$ with $\beta = 2.5\alpha$. In Fig. 5.4 we plot the cumulative number counts per square degree of sources of given flux in the 2QZ radio-loud sample. We determine the slope of the number-counts, $\alpha$ (equivalent to the slope of the power-law form for magnitudes) by interpolating between adjacent bins in logarithmic space to work out $\beta$ (the flux power law slope) and converting to $\alpha$. Our resulting values of the slope agree well with those derived by Condon et al. (1998) for all objects in the NVSS, who found $\alpha = 0.4$ at around 35mJy flux.

Two dashed lines are marked in Fig. 5.4, at $\alpha = 0.3$, which is the faint-end optical number-counts slope and at $\alpha = 0.4$, the value of the slope for which no lensing is expected. Several things can be taken from this plot. The slope is very shallow ($\alpha < 0.1$) for radio-loud QSOs fainter than about 7mJy (there are around 100 such QSOs in this sample). For a magnification of 1.5 (the magnification typical for a slope of $\alpha = 0.3$ with enhancement factor 0.9), a slope of 0.1 would yield an enhancement factor of only 0.7. This means we might expect large deficits due to lensing at low radio fluxes (although we approach the completeness threshold of the NVSS at 2.5 mJy, so the slope will become artificially flat). However, this would not be the case when we take into account Equation 5.4. Due to double magnification bias, the expected slope for the lensing effect in a radio-and-optically selected QSO subsample is the sum of the radio and optical slopes. As the optical number-count slope for 2QZ QSOs is never less than about 0.3, the actual lensing slope for the faintest radio sample is near 0.4, where no lensing effect is expected. This illustrates an even more important point for our purposes. As the radio counts never have a negative slope, the sum of the radio and optical slopes for the 2QZ sample can *never* be less than the optical slope. As we have seen in Chapter 3, if the value of $\alpha$ lowers, the QSO-galaxy correlation function decreases. This means that, *under the assumption that the radio and optical counts are independent*, the radio-optical double magnification bias could never enhance an anti-correlation between QSOs and galaxies.

It is worth noting that the simple double magnification bias *can* enhance the strength of the signal when a positive correlation is expected. In fact, the double magnification bias in radio-and-optically selected QSO subsamples will *always* somewhat increase positive correlations between QSOs and galaxies. Fig. 5.4 shows that the counts slope approaches 0.75 for QSOs with radio fluxes larger than about 100mJy. As the optically bright QSO slope is also around 0.75, the enhancement factor will more than double (for a magnification of 1.5, as above) in a sample of QSOs that is bright in both radio and optical bands. Although such a sample of QSOs only comprises about 1 per cent of the 2QZ, there are

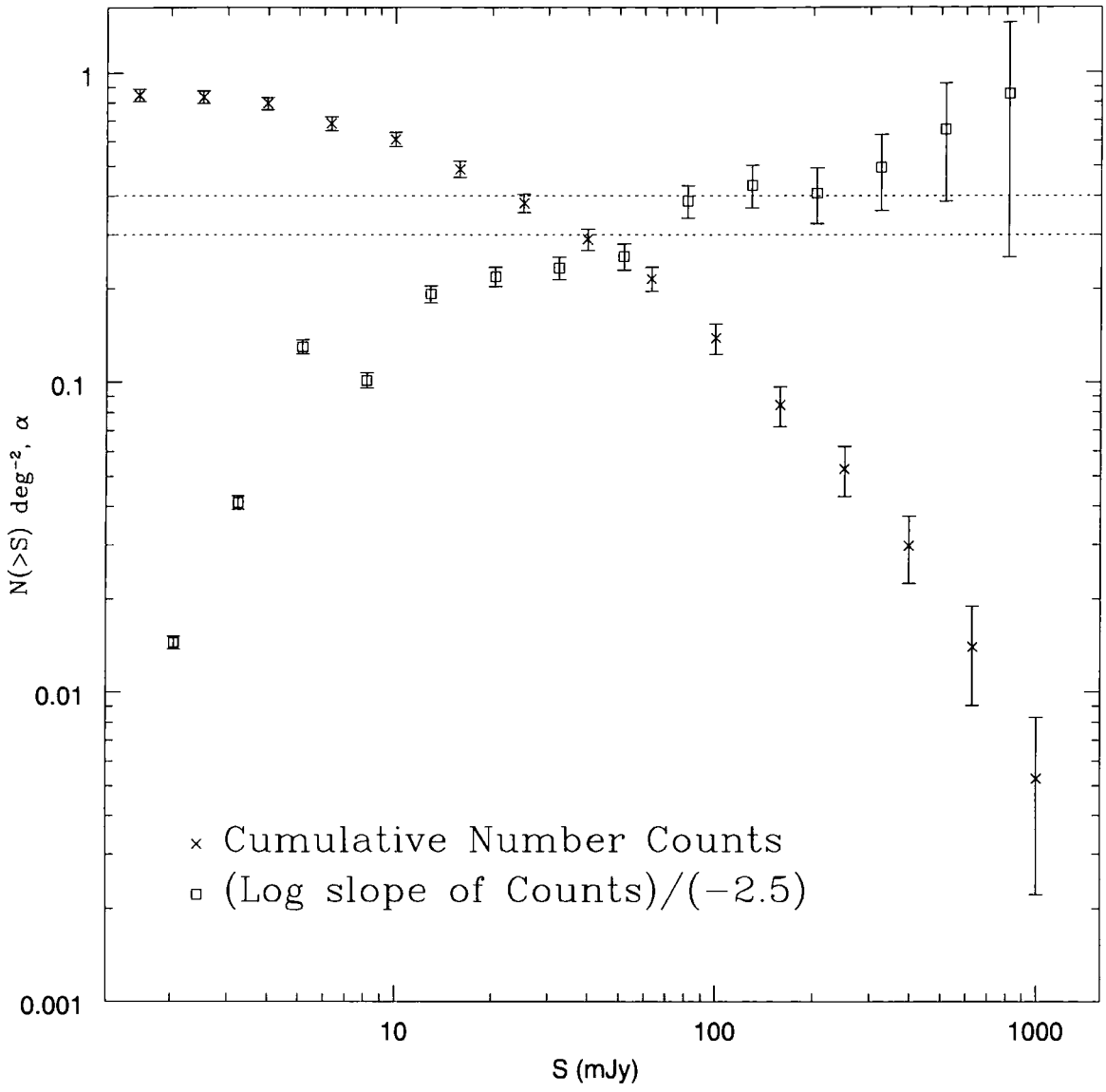


Figure 5.4: This figure shows the cumulative number-counts (per square degree) by flux of radio-loud QSOs in the 2QZ as crosses, with Poisson error bars. The squares mark the value of the slope, $\alpha$, of the counts, determined by interpolating between adjacent bins of the number-counts. Dashed lines are drawn at $\alpha = 0.3$, corresponding to the typical faint-end slope of the optical number-counts, and $\alpha = 0.4$, the slope value for which no statistical lensing effect is expected.

326 QSOs with radio-fluxes (out of a total of 3814) in the SDSS EDR QSO catalogue (Schneider et al. 2002), meaning there should be more than 8000 radio-loud QSOs in the final SDSS catalogue. The SDSS radio-loud sample should ultimately be very useful in studying QSO-galaxy cross-correlations, although care must be taken when looking for the effects of double magnification bias, as the SDSS QSO sample has a complicated radio/optical selection procedure (Richards et al. 2002) and one should be circumspect in preparing a QSO subsample selected in *both* radio and optical bands rather than in *either* band.

Much of this section has introduced double magnification bias under the assumption that the distribution of radio and optical fluxes of the considered QSO subsample are independent. Wyithe et al. (2003) have extended the discussion of double magnification bias arising from QSO samples that are correlated in different passbands. In Fig. 5.5 we show $b_J$ magnitude plotted against radio flux for the 2QZ radio-loud sample. Although the distribution is highly scattered and looks fairly random, the two passbands are in fact highly correlated (at better than the 99 per cent significance level), having a Spearman rank correlation co-efficient of -0.154. In samples where passbands are correlated but with a great deal of scatter, Wyithe et al. (2003) argue that the expected enhancement due to magnification bias can be greatly altered in counterintuitive ways, that could, perhaps, lead to significant anti-correlations between QSOs and foreground lenses. Wyithe et al. (2003) further argue that such an alteration could be even more pronounced if the correlation between the passbands is non-linear. In Section 5.4, we will check the cross-correlation between the radio-loud 2QZ subsample and galaxies, to search for any strong signal that would be hard to attribute to anything other than a lensing effect.

## 5.4 QSO and Galaxy Cross-Correlation Functions

The two-point angular correlation function (Peebles 1973, Hauser & Peebles 1973) quantifies the joint probability of detecting two sources, separated by a given angular distance, within a given solid angle (see Appendix B). It is the main statistic of choice in studies of how QSO and galaxy distributions are related, although how the statistic is estimated varies considerably. In this section, we measure the cross-correlation between SDSS or APM galaxies and 2QZ (or 6QZ) QSOs. We study whether the signal may arise as a selection effect and consider different explanations for the signal. Notably, the expected evolution of the QSO-galaxy cross-correlation with redshift and (especially) magnitude

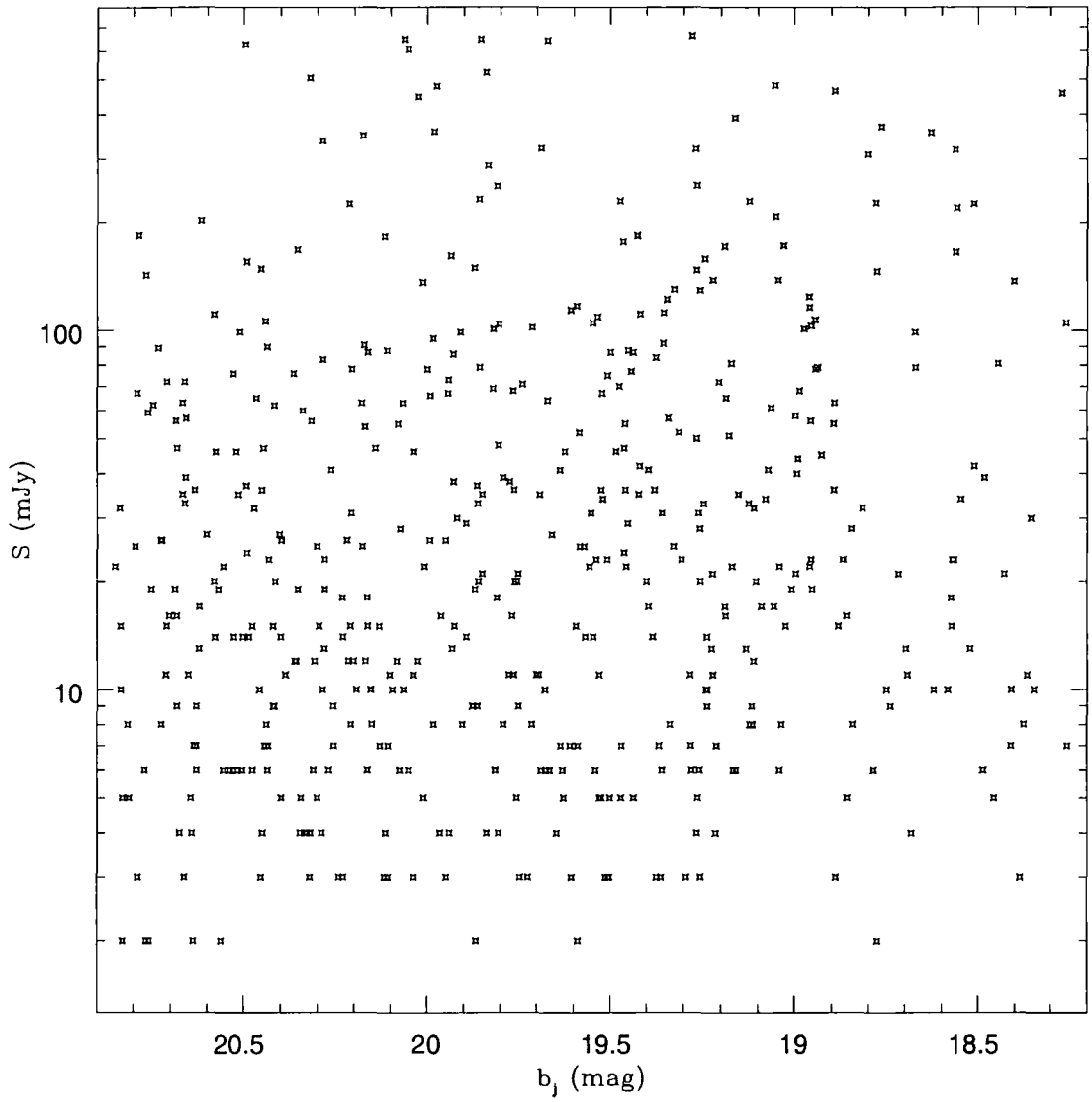


Figure 5.5: This plot illustrates the scatter of radio-flux against $b_J$ magnitude for the 2QZ radio-loud sample. Although the distributions in the two passbands seem fairly random, they are actually significantly correlated, with a Spearman rank correlation co-efficient of -0.154. This could be important, as Wyithe et al. (2003) have determined that subsamples of QSOs selected in bands that are correlated with a great deal of scatter can result from counterintuitive magnification bias. Such subsamples may show enhanced lensing effects.

of the QSO sample is a strong prediction of the statistical lensing hypothesis, and is something we may be able to test with the 2QZ sample.

### 5.4.1 Correlation Estimator and Errors

To measure the two-point correlation function, $\omega(\theta)$, we use the estimator derived by Peebles & Hauser (1974), as discussed in previous chapters and defined in Equation 3.5. As there are far more sources in the galaxy samples used in this chapter then in the group catalogues considered in previous chapters, we produce random catalogues with only 10 objects for every data point. The random catalogues have the same angular selection function as the considered data. As in previous chapters, we carry out the correlation analysis in 30 contiguous areas smaller than the full 2QZ survey, which correspond to plate boundaries in the APM. This process acts to reduce correlations that may arise from variable magnitude limits on individual plates across the APM survey. The 2QZ QSO sample we shall use throughout this chapter is as outlined in Section 4.3.3 of the previous chapter - the subsample of 2QZ QSOs with "11" quality spectroscopic identifications and redshifts greater than 0.4.

In previous chapters we have introduced several error estimates on the correlation function, *Poisson Error*, *Expected Poisson Error*, *Field-to-Field Error* and *Jackknife Error*. We also outlined the method by which we calculate the *covariance* from Monte Carlo simulations of the data. We shall now briefly test these error estimates by creating 100 Monte Carlo simulations with the same size and angular selection as the 2QZ QSO sample, cross-correlating them against the SDSS and APM galaxy samples, and compare the various error estimates to $1\sigma$ standard deviations in the values of $\omega(\theta)$ for the Monte Carlo simulations.

In the lower panel of Fig. 5.6, we display the mean cross-correlation signal across the 100 Monte Carlo simulations for the SGC QSOs (cross-correlated against APM galaxies), the NGC QSOs (cross-correlated against SDSS EDR galaxies) and the combined result for the two strips. The agreement between the NGC and SGC results is excellent - better than 12 per cent of the Monte Carlo error on the NGC mean over all scales. The deviation of the combined result from zero, the expected result for random samples, is similarly no more than 12 per cent of the combined Monte Carlo error over all scales. We note that the shot noise would comprise 10 per cent of the Monte Carlo error (as we have 100 samples). Although the correlation function seems to begin to diverge on scales smaller than 0.3 arcminutes, the error is sufficiently large on these scales that we may consider

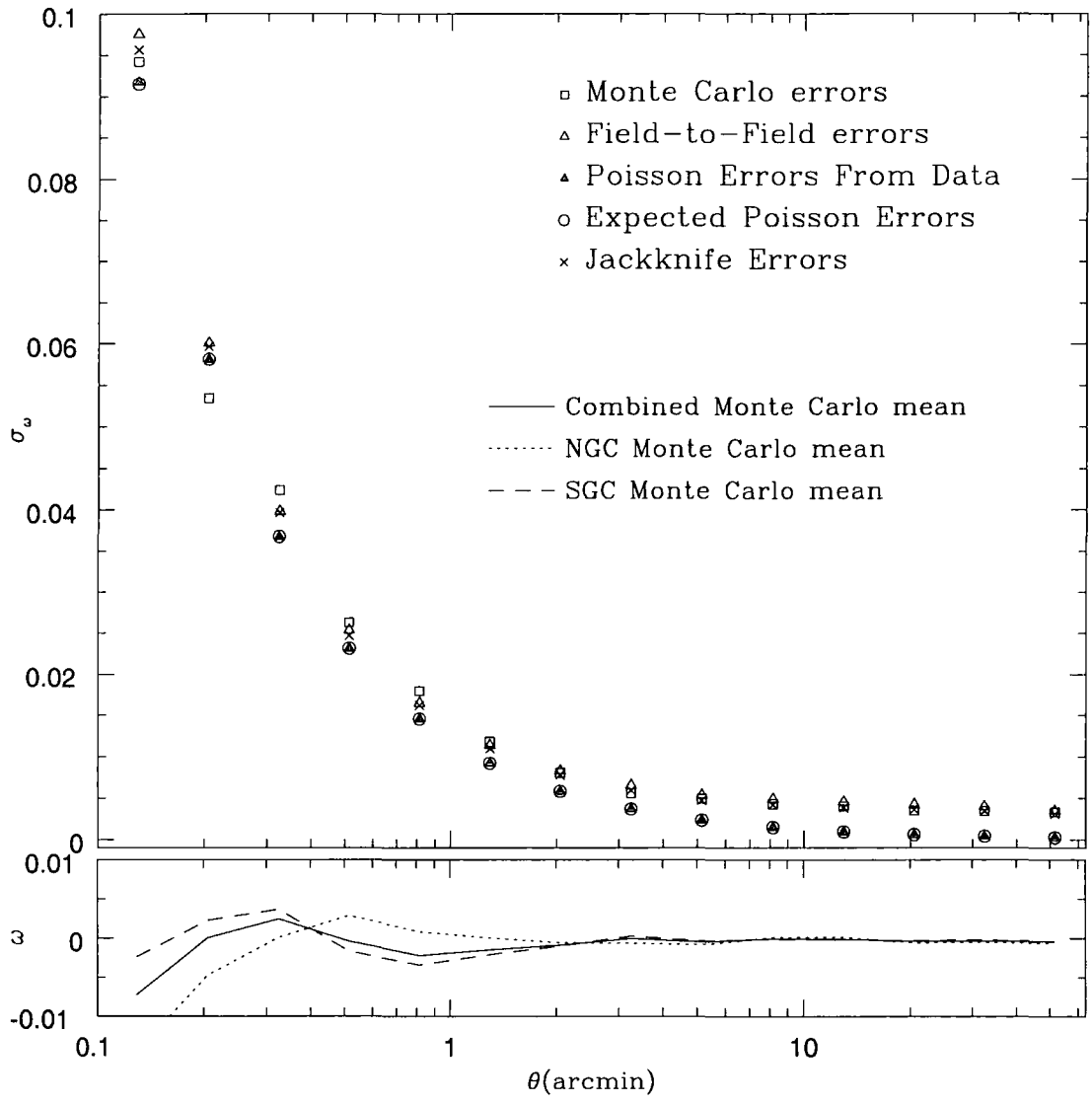


Figure 5.6: In the upper panel, we display error estimates on the correlation function for each of the equations mentioned in the text. The plotted estimates are the averages from 100 Monte Carlo simulations combined for the NGC and SGC strips. Also plotted are the Monte Carlo errors - the standard deviation of the 100 simulations. In the lower panel we plot the mean of the 100 simulations for the NGC and SGC individually and for the NGC and SGC combined.

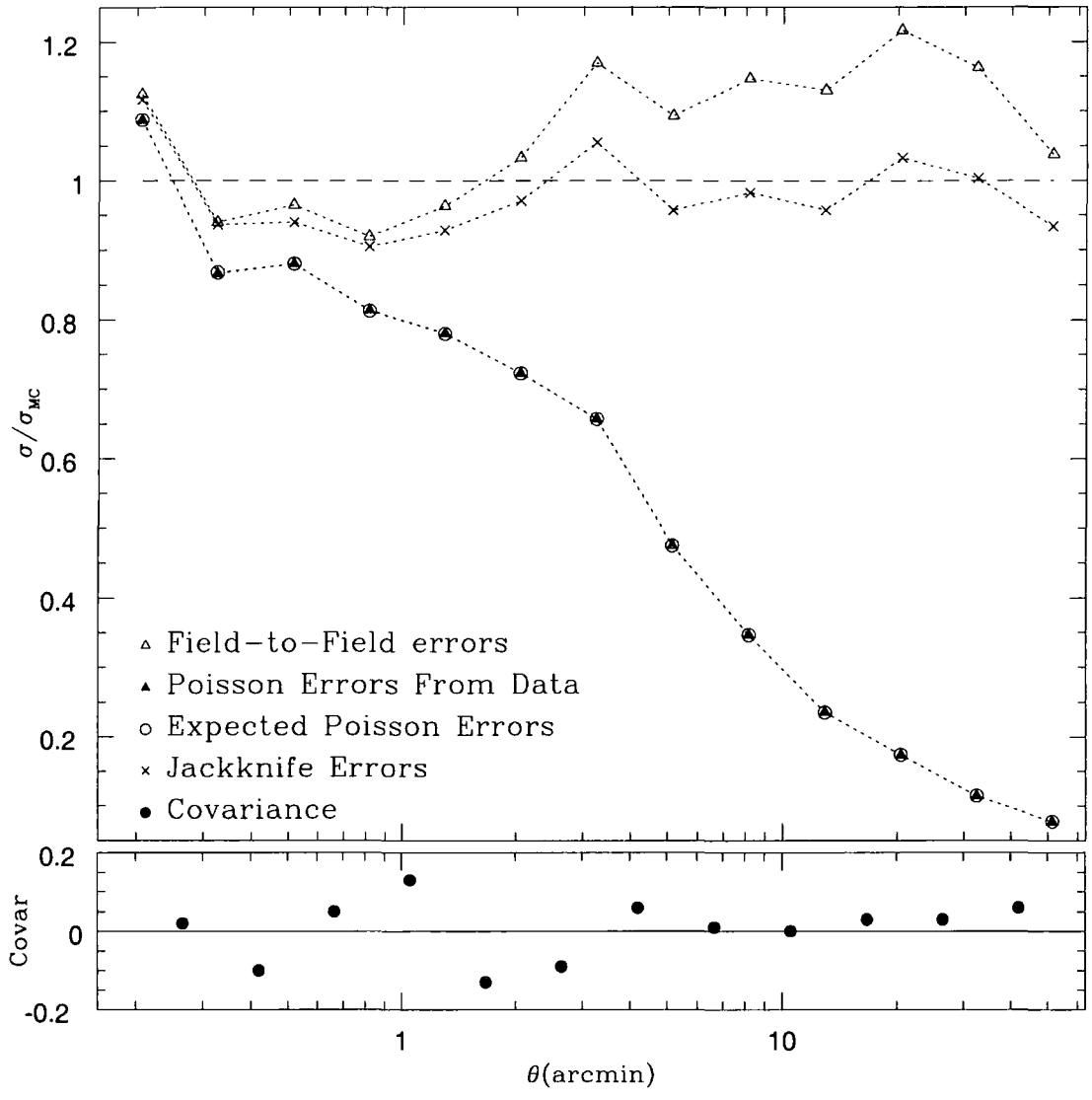


Figure 5.7: In the upper panel we display the error on the correlation function taken in ratio to the Monte Carlo error estimate, for each of the equations mentioned in the text. In all cases the errors are determined for the combined NGC and SGC sample. The dashed line at $\sigma/\sigma_{MC} = 1$ (where the Monte Carlo estimate itself would lie) is drawn for comparison. The lower panel depicts the covariance between adjacent bins determined from 100 Monte Carlo realisations.

the correlation estimator probably valid on all of the scales plotted and certainly valid on scales larger than 0.3 arcminutes. Note that the consistency of the correlation estimator across all scales indicates that the software we use to calculate the estimator is robust.

In the upper panel of Fig. 5.6 we plot the various error estimates outlined in this section. The general trend of the errors is in good agreement, although the Poisson error estimates begin to under-predict the value of the error (as compared to the Monte Carlo estimate) on larger scales. We assume that the Monte Carlo error represents a fair estimate of the true error on the correlation function. In Fig. 5.7 we plot the various errors taken in ratio to the Monte Carlo error. It is obvious that the Poisson error is an underestimate on even quite small scales and underestimates the error by as much as 50 per cent even on arcminute scales. The jackknife and field-to-field error estimates are much better and both constitute reasonable estimates of the error on scales from 0.2 arcminutes to nearly a degree. The field-to-field error, however, is perhaps a 20 per cent overestimate on larger scales, where the jackknife error remains in line with the Monte Carlo estimate. In Fig. 5.7, we also plot the covariance between *adjacent* bins. The covariance is quite low - almost within the 10 per cent expected standard error arising from the 100 Monte Carlo simulations. Throughout the remainder of this section, we adopt the estimate of the correlation function defined by Equation 3.5, together with the jackknife error estimate of Equation 4.5 as both seem fair over the scales that we will probe. We will often quote the significance of results by estimating the correlation function and it's associated error in one "large" bin (usually out to 10 arcminutes), an approach that makes the, already quite low, covariance negligible.

### 5.4.2 The Cross-Correlation of 2QZ QSOs and Galaxies

In Fig. 5.8 we plot the cross-correlation of all 2QZ QSOs that meet our usual selection criteria ( $z > 0.4$ and 2QZ identification of "11") against SDSS EDR galaxies (in the region of the 2QZ NGC strip) and APM galaxies (in the region of the 2QZ SGC strip). The upper panel of Fig. 5.8 shows the cross-correlation individually for the NGC and SGC strips. Galaxies are anti-correlated with QSOs in both strips. The anti-correlation is slightly stronger in the NGC strip but not significantly so. In the lower panel of Fig. 5.8, we plot the cross-correlation for both "directions". A significant anti-correlation is detected irrespective of whether we centre on galaxies and count QSOs or centre on QSOs and count galaxies, indicating that our angular completeness masks consistently account for the angular selection of QSOs or galaxies.

If we bin the data displayed in Fig. 5.8 in a single bin of extent 10 arcminutes and estimate the correlation function and ($1\sigma$) jackknife error, we find there is an anti-correlation of strength $\omega(< 10') = -0.007$ out to 10 arcminutes (about 1 $h^{-1}$Mpc at the median galaxy redshift of 0.15). The significance of this result is $2.8\sigma$ for $\omega_{qg}$ and $2.2\sigma$ for $\omega_{gq}$. Although the anti-correlation is slightly less significant for $\omega_{gq}$, it is also slightly stronger; $\omega(< 10') = -0.008$. It is unclear why the errors are slightly larger when the analysis is carried out centring on galaxies but it may simply be due to the fact that there are many more galaxies in the samples used than QSOs, so the background density of objects is better constrained.

The anti-correlation displayed in Fig. 5.8 becomes very strong on small scales and agreement between the two "directions" of the correlation function is excellent. For instance, both $\omega_{qg}$ and $\omega_{gq}$ show an anti-correlation of strength -0.02 out to 3 arcminutes at $3\sigma$ significance. Note that the innermost bin barely contributes to this particular signal, as it contains less than one-hundredth of the pairs in the bin at 2 arcminutes. In following sections, we will often display $\omega_{qg}$, the slightly weaker, slightly more significant result. This is mainly because ultimately, when discussing the anti-correlation in terms of lensing, we will compare the galaxy-galaxy auto-correlation, $\omega_{gg}$, to $\omega_{qg}$, which shares the same random catalogue. It is useful, though, to have shown that there is a signifcant, consistent anti-correlation between QSOs and galaxies irrespective of which masks we use to compute the correlation function.

### 5.4.3 Is the Anti-Correlation Between Galaxies and QSOs a Selection Effect?

Certainly, there is a significant anti-correlation between galaxies and QSOs. It is natural to ask whether the signal arises during the process of constructing the QSO or galaxy catalogues. There are several obvious measurement procedures that might produce an anti-correlation between QSOs and galaxies. The initial construction of the 2QZ UVX target catalogue involved removing all objects that weren't stellar in appearance. Although all "high" redshift ($z \gtrsim 0.5$) QSOs should be stellar in appearance, they may merge with foreground objects creating objects that look extended on images. A bright ($b_J < 19.5$) subsample of these extended images should end up in the 2dF Galaxy Redshift Survey (henceforth 2dFGRS) and thus appear in deficit in the 2QZ. It turns out that this influences separations between galaxies and QSOs on scales of about 8 arcseconds (Madgwick et al. 2002), much smaller than the scales we are probing. Restrictions on

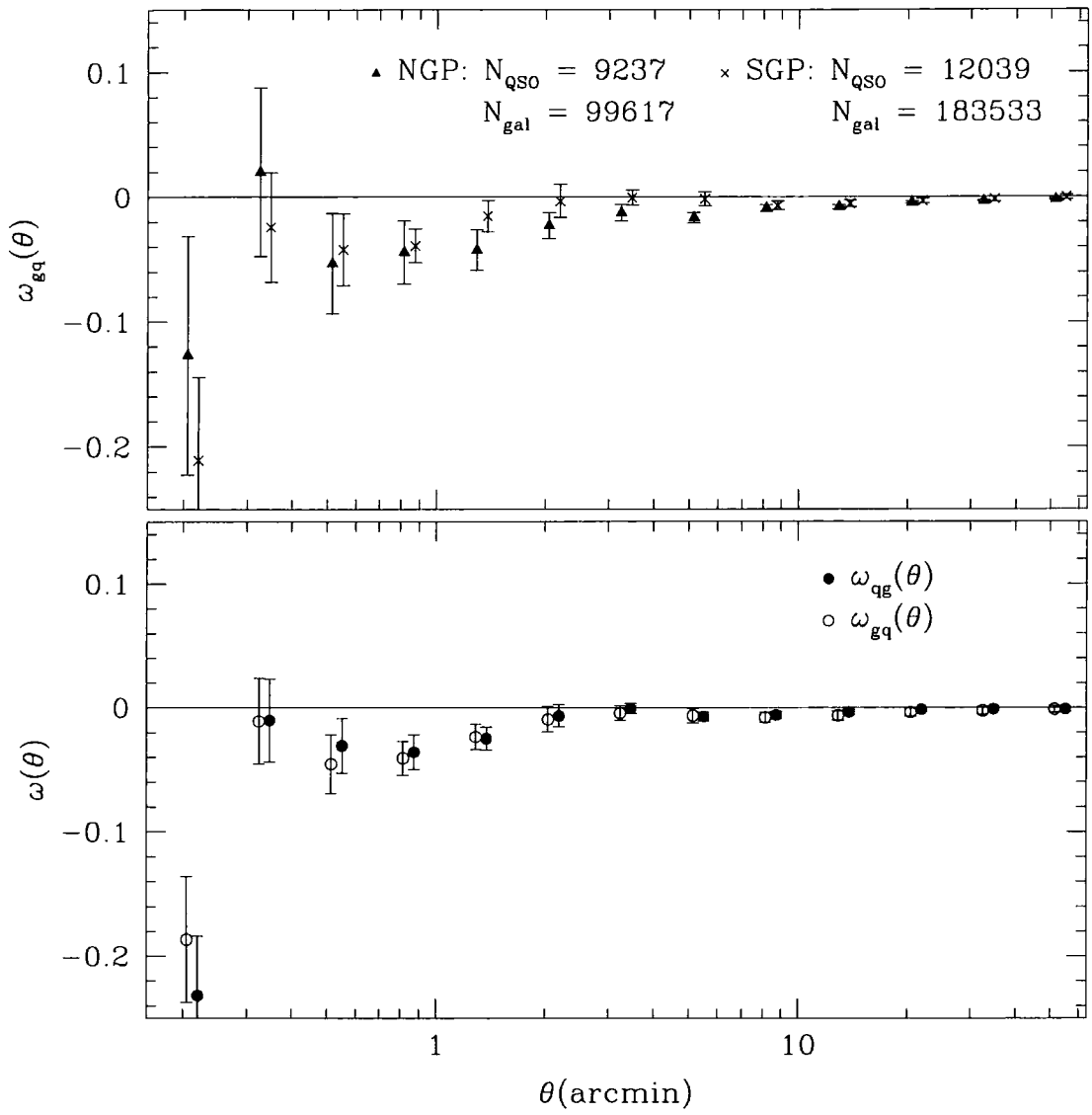


Figure 5.8: This figure plots the cross-correlation of all QSOs in the 2QZ against SDSS EDR galaxies (in the region of the NGC 2QZ strip) and against APM galaxies (in the region of the SGC 2QZ strip). The upper panel displays the cross-correlation signal for the 2 strips individually. The lower panel displays the signal combined for both strips. The lower panel shows estimates for both "directions", centring on QSOs and counting galaxies ($\omega_{qg}$) and centring on galaxies and counting QSOs ($\omega_{gq}$). Error bars represent $1\sigma$ jackknife errors. Labels note the number of objects of each population present within the confines of the 2QZ boundaries. Points within the same bin have been offset slightly for ease of display.

the placement of 2dF fibres means that the minimum angle between objects in any 2QZ field is about 30 arcseconds, which might mean a paucity of objects at small separations. This restriction would include QSO-galaxy separations as the 2QZ was carried out simultaneously with the 2dFGRS. However, the vast majority of fields in the 2dF survey were observed several times to overcome this problem and we have already suggested in Chapter 3 that restrictions on fibre separation have a minimal effect on 2QZ populations.

The easiest way to judge measurement systematics in the 2QZ is to consider a control sample of objects that underwent identical data reduction as the QSOs but should display no cosmological signatures. There are 10,587 stars (with "11" identification quality) in the 2QZ catalogue. In Fig. 5.9 we plot the cross-correlation of these stars against our galaxy samples. The upper panel of Fig. 5.9 compares the cross-correlation estimate for the NGC and SGC 2QZ strips. The agreement is reasonable, although the NGC sample is slightly more positively correlated with galaxies across all scales. Note that we might not expect the stellar correlation functions to be zero on all scales - gradients exist in the stellar distribution as we pan across our Galaxy and these gradients are not recreated in the random catalogues. Such gradients may make the correlation signal higher or lower on average. However, we would expect the stellar signature to be *flat* across the scales of interest. This is highlighted in the lower panel of Fig. 5.9, where we display the star-galaxy and galaxy-star cross correlations. Unlike in the QSO-galaxy case, there is a discrepancy in the large-scale zero-point of the correlation function that depends upon the "direction" of the cross-correlation. When the random catalogue is constructed to match the stellar distribution, the zero-point of the correlation function drops significantly (to about -0.03). The large-scale value of the correlation function *is* zero when the random catalogue is constructed to match the galaxy distribution, which is free from (at least genuine physical) gradients. The key point, is that the correlation functions for stars and galaxies are flat across our scales of interest (deviating at most $0.5\sigma$ from their respective zero-points on scales larger than $\theta > 0.4$ arcminutes), indicating that systematics in the construction of the galaxy and QSO samples are low, hence induce no false correlations in our QSO-galaxy cross-correlations. The innermost points ($\theta < 0.4$ arcminutes) plotted in Fig. 5.9 seem to genuinely deviate from the zero-point and may be representative of merged images, fibre placement signatures or some other systematic. We will continue to plot these points in figures in this section but will not consider them in any modelling analysis or quotes of the significance of a signal.

For completeness, we display the cross-correlation between our galaxy samples and

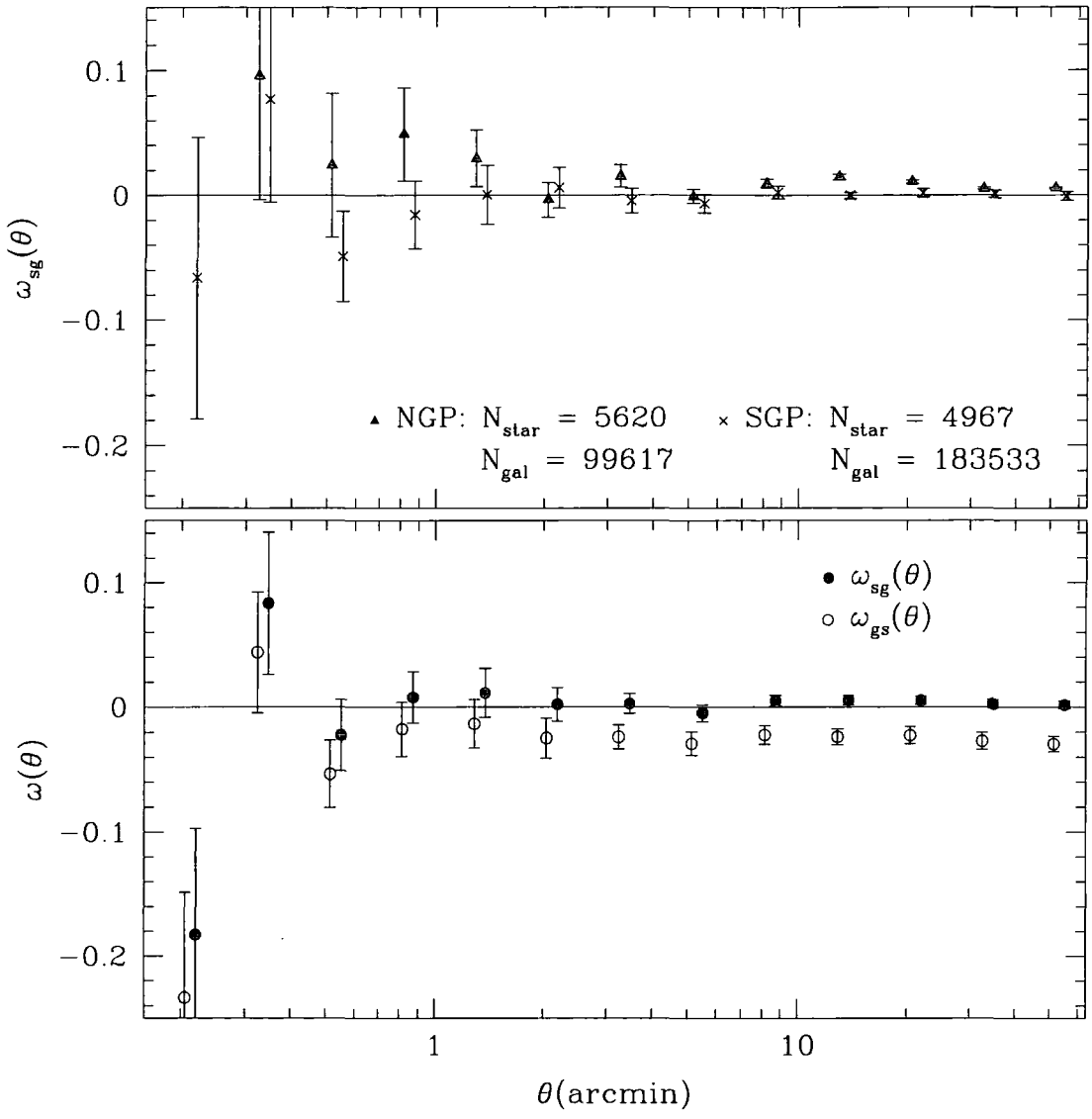


Figure 5.9: This figure plots the cross-correlation of all stars in the 2QZ against SDSS EDR galaxies (in the region of the NGC 2QZ strip) and against APM galaxies (in the region of the SGC 2QZ strip). The upper panel displays the cross-correlation signal for the 2 strips individually. The lower panel displays the signal combined for both strips. The lower panel shows estimates for both "directions", centring on stars and counting galaxies ($\omega_{sg}$) and centring on galaxies and counting stars ($\omega_{gs}$). Error bars represent $1\sigma$ jackknife errors. Labels note the number of objects of each population present within the confines of the 2QZ boundaries. Points within the same bin have been offset slightly for ease of display.

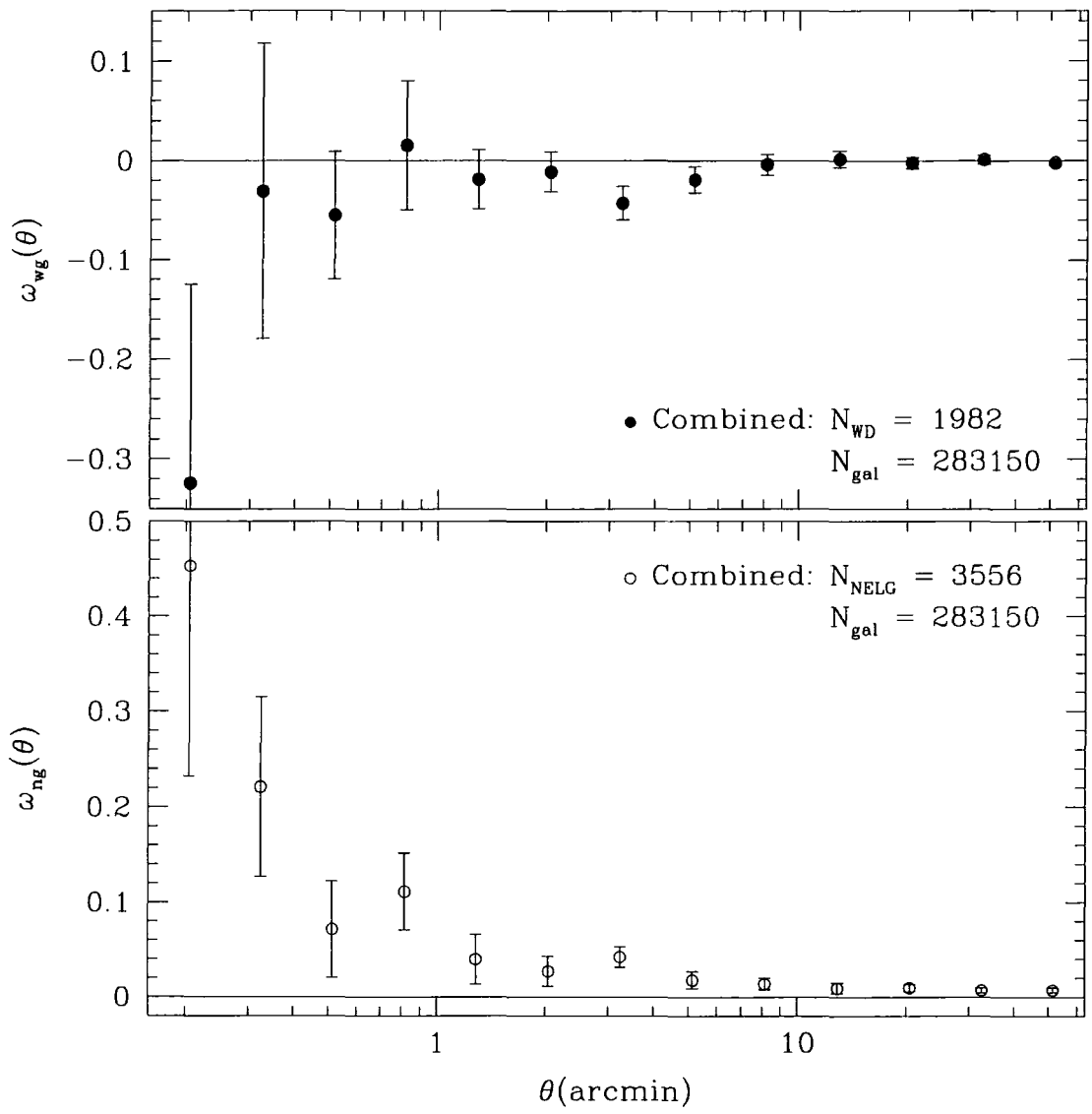


Figure 5.10: This figure plots the cross-correlation of some dominant populations of objects in the 2QZ against SDSS EDR galaxies (in the region of the NGC 2QZ strip) and against APM galaxies (in the region of the SGC 2QZ strip). The upper panel displays the cross-correlation signal for 2QZ White Dwarfs against galaxies (for the NGC and SGC strips combined). The lower panel displays the cross-correlation signal for 2QZ Narrow Emission Line Galaxies against galaxies (for the NGC and SGC strips combined). Error bars represent $1\sigma$ jackknife errors. Labels note the number of objects of each population present within the confines of the 2QZ boundaries. Points within the same bin have been offset slightly for ease of display.

other dominant samples of objects drawn from the 2QZ, White Dwarfs and Narrow Emission Line Galaxies (henceforth NELGs), in Fig. 5.10. The upper panel displays the cross-correlation of 2QZ White Dwarfs and galaxies. Tantalisingly, there is a $2.0\sigma$ anti-correlation on 3-7 arcminute scales. We have seen an anti-correlation on these scales before in Chapter 3 and given that previous results in this section suggest no measurement systematics in our correlation estimator, it is worth considering if this slight but significant result could arise from a physical process. The obvious mechanism that would cause a deficit of galaxies around White Dwarfs is circumstellar dust distributed on average in a 3-7 arcminute shell around the star. The average apparent magnitude of the sample of White Dwarfs displayed in Fig. 5.10 is $b_J = 19.8$. A fair median value for the absolute magnitude of a White Dwarf is about 13.5 (Sion 1984, Liebert et al. 1988), which would place our White Dwarf distribution at an average distance of 180pc. At this distance, 3-7 arcminutes translates into 0.15-0.35pc or about 30,000 to 70,000 AU from the central star, where one might expect to find Oort cometary cloud systems (Oort 1950, Hills 1981). White Dwarf systems have many similarities to our own, and Oort clouds around White Dwarfs have been proposed indirectly as an explanation for dust effects on scales much closer to White Dwarfs (Debes & Sigurdsson 1997). The idea that we have detected cometary clouds around White Dwarfs is, of course, highly speculative but is certainly worthwhile considering if the effect persists with larger significance in larger samples of White Dwarfs.

In the lower panel of Fig. 5.10, we display the cross-correlation of NELGS against our galaxy samples. As might be expected due to general clustering of galaxes (65 per cent of NELGS in the 2QZ are at redshifts of $z < 0.3$), 2QZ NELGS are strongly positively correlated with other galaxy samples.

The main point of this subsection has been to show that there are no measurement systematics that might contaminate the significant anti-correlation between 2QZ QSOs and galaxies. We will now consider cosmological explanantions for the anti-correlation.

### 5.4.4 Cosmological Explanations For the Anti-Correlation Between QSOs and Galaxies

An obvious physical effect, other than statistical lensing, that might be proposed as the cause of the anti-correlation between QSOs and galaxies is dust in galaxies obscuring background QSOs. This would lead to a dearth of QSOs around galaxies. Alternatively, If the anti-correlation *is* due to statistical lensing, we might be able to see the effects of

double magnification bias (see Section 5.3, above) in the radio-loud 2QZ sample. A second signature of statistical lensing we can check for is the evolution of the cross-correlation between galaxies and QSOs with the magnitude of the QSO sample.

#### 5.4.4.1 Is Intervening Dust the Cause of the Anti-Correlation Between QSOs and Galaxies?

We can use colour information in the 2QZ to determine if intervening dust preferentially distributed around galaxies could remove QSOs from the 2QZ catalogue out to 10 arcminutes ($\sim 1\ h^{-1}$Mpc). Our method is similar to the correlation estimator of Equation 3.5 but instead of counting the average *number* of QSOs in differential annuli around galaxies, we work out the average *colour* of QSOs in the same annuli. To check what the expected result for this procedure is, we perform a bootstrap analysis of the colours of QSOs around galaxies. We have created 1000 "scrambled" QSO catalogues that have the exact same positions and colours as the 2QZ sample but the colours are randomly matched to the positions. In essence, each QSO in the 2QZ has been randomly reassigned a new colour from the distribution of 2QZ QSOs. As is the norm for bootstrap analysis, the colours are selected with replacement to ensure a random element to the "scrambled" catalogues (see, for instance, Barrow, Bhavsar & Sonada 1984). The average colour of QSOs in annuli around galaxies is calculated for the scrambled QSO catalogues as for the real data and the median and ($1\sigma$) deviation across the 1000 scrambled results is determined and quoted as our expected value and errorbar.

We can place limits on the amount of dust around galaxies along QSO lines of sight by considering the range of the error bars in any bin. Knowing the expected value on degree scales and the range of allowed measured values in each bin, we can calculate the allowed colour excess in a given bin for both sets of measured 2QZ colours, $u - b_J$ and $b_J - r$. The colour excess, $E(B-V)$, is defined

$$E(B-V) = (B-V)_{OBS} - (B-V)_{TRUE} \tag{5.5}$$

for $B-V$ colours, where "OBS" denotes the observed colour and "TRUE" the expected colour. Schlegel, Finkbeiner & Davis (1998) provide tables to convert from the colour excess to the amount of absorption by dust in magnitudes for many passbands. Schlegel, Finkbeiner & Davis (1998) base their absorption laws on the difference in reddening between local, lightly-reddened standard stars and more distant reddened stars in the

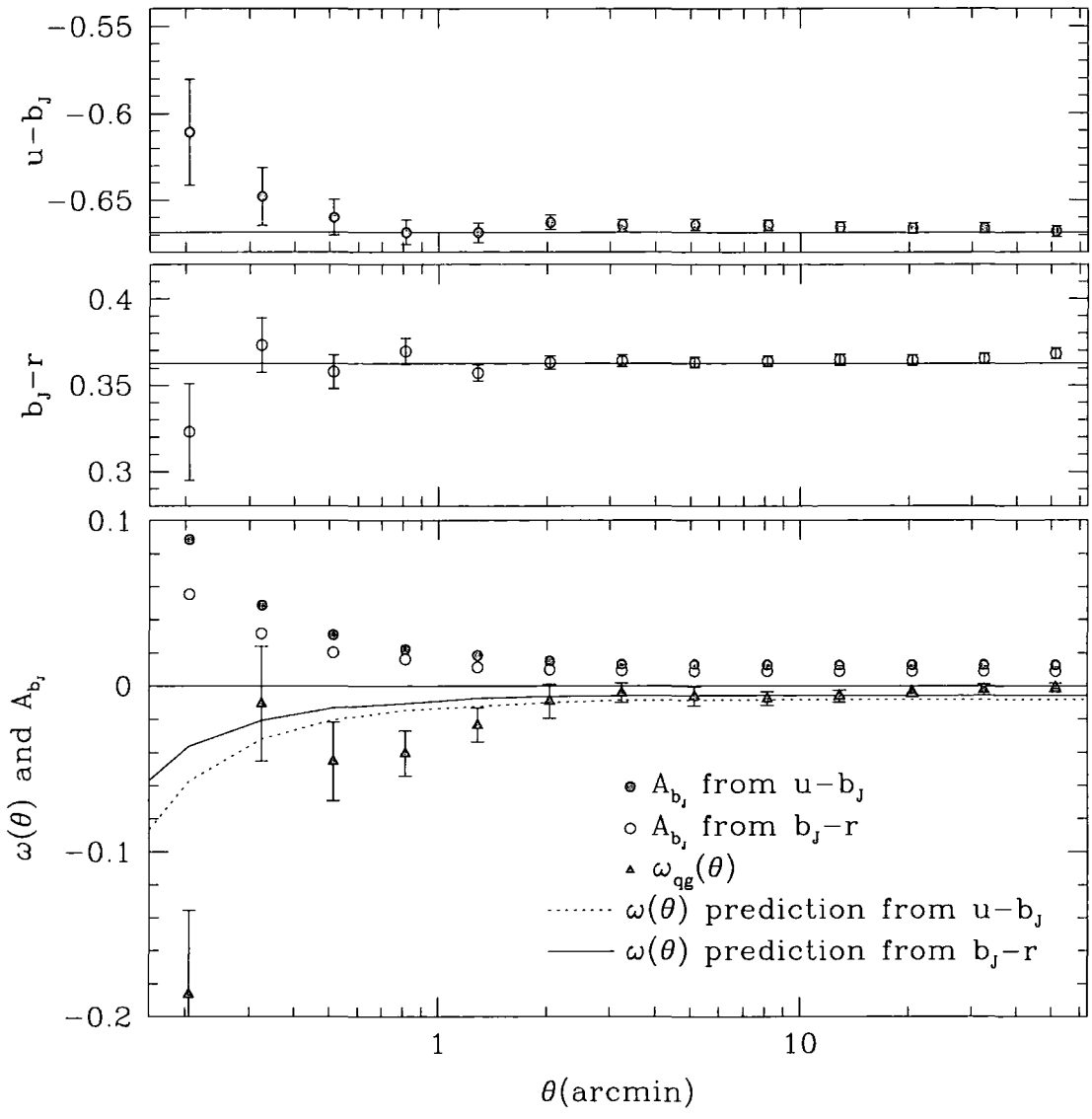


Figure 5.11: This figure places limits on the amount of "typical" dust that could account for the QSO-galaxy anti-correlation meeasured in Fig. 5.8. The upper two panels show the average colour of QSOs in bins centred on SDSS EDR galaxies (in the region of the NGC 2QZ strip) and against APM galaxies (in the region of the SGC 2QZ strip). Both the results for $u - b_J$ and $b_J - r$ colours are shown. Error bars in these panels represent the standard deviation ($1\sigma$) in 1000 bootstrapped simulations with the same QSO positions as the 2QZ catalogue but scrambled colours. The lower panels translate these $1\sigma$ error bars into limits on absorbtion in the $b_J$ band ($A_{b_J}$). The absorption limits are translated into limits on the QSO-galaxy cross-correlation using a simple model outlined in the text. These limits are then displayed against the points and errors on $\omega_{qg}$ from Fig. 5.8, represented by triangles.

Milky Way (Cardelli, Clayton & Mathis 1989, O'Donnell 1994). The reddened stars are chosen to sample a wide range of interstellar environments (Fitzpatrick & Massa 1990) and Schlegel, Finkbeiner & Davis (1998) demonstrate that their dust laws excellently reproduce the reddening (as compared to the MgII index) of a sample of nearly 500 elliptical galaxies that have broad sky coverage. However, by design, the absorption law used by Schlegel, Finkbeiner & Davis (1998) only applies to the Milky Way Galaxy. There is no reason to believe that other galaxies have the same dust laws as our own. In fact, although the Magellanic Clouds have been shown to have similar absorption laws to the Milky Way (Koornneef 1982, Bouchet et al. 1985), other local ($z \lesssim 0.03$) galaxies have been shown to have "greyer" absorption laws (Calzetti, Kinney & Storchi-Bergmann 1994, Kinney et al. 1994), meaning that more dust absorption is expected for the same amount of reddening (at optical wavelengths). When quoting results based on the dust laws used by Schlegel, Finkbeiner & Davis (1998), we will often add the caveat that we are assuming "typical" dust, by which we mean a dust law that is typical for the Milky Way Galaxy, rather than the greyer law derived by Calzetti, Kinney & Storchi-Bergmann (1994). We can demonstrate that the Calzetti, Kinney & Storchi-Bergmann (1994) absorption law is not significantly different to the Milky Way absorption law, for our purposes. The $B - V$ colour excess may be characterised by the equation

$$E(B-V) = A_B - A_V = \frac{A_V}{R_V} \tag{5.6}$$

(Schlegel, Finkbeiner & Davis 1998), where $R_V$ is a function of the particular absorption law studied. Bolzonella, Miralles & Pelló (2000) derive the values $R_V = 3.1$ for the Milky Way (see also Schlegel, Finkbeiner & Davis 1998) and Large Magellanic Cloud, $R_V = 2.72 \pm 0.21$ for the Small Magellanic Cloud and $R_V = 4.05 \pm 0.80$ for a Calzetti, Kinney & Storchi-Bergmann (1994) absorption law. Given that $A_{b_J} = 0.72A_B + 0.28A_V$ (Schlegel, Finkbeiner & Davis 1998) it is straightforward to show that assuming a Milky Way absorption law underestimates the absorption in the $b_J$ band by only $\sim 25$ per cent compared to the greyer Calzetti, Kinney & Storchi-Bergmann (1994) law.

We convert our colour excesses into absorption assuming the Milky Way dust law used by Schlegel, Finkbeiner & Davis (1998). We then convert our $1\sigma$ limits on absorption into a limit on the correlation function using a simple model outlined in Boyle, Fong & Shanks (1988). The correlation function, $\omega$ is equivalent to one subtracted from the enhancement factor, $q$. Dust around galaxies will cause an absorption of a given number of magnitudes (which we calculate from our measured allowed colour excesses) that will

alter the magnitude limit of the QSO number-counts close to galaxies

$$\omega(\theta) + 1 = q = \frac{N(< m)_{Gal}(\theta)}{N(< m)_{Field}} = \frac{N(< m)_{Field} - A_{b_J}(\theta)}{N(< m)_{Field}} \tag{5.7}$$

where $N(< m)$ represents the integrated number-counts by magnitude, $A_{b_J}$ is the absorption in the $b_J$ band and the subscripts "Gal" and "Field" represent values near to galaxies and in the field, respectively. Equation 5.5 can easily be simplified if the number-counts are represented by a power law (c.f. Equation 3.13) but to be exact, we will simply use the full fitted (SPL) form of the number-counts of 2QZ QSOs displayed in Chapter 2 above.

We now have a simple model that converts the error on our measurement of the average colour of QSOs near galaxies into a limit on the observed anti-correlation due to dust. In the upper two panels of Fig. 5.11, we show the average colour of QSOs around our galaxy samples with bootstrapped error bars, for both colours available in the 2QZ catalogue. A solid line marks the expected value from our scrambled simulations. Except, perhaps, for the two innermost bins, which we have mentioned in Section 5.4.3 could arguably be discarded from our analysis, there is no significant deviation in the colour of QSOs from the expected value. Note that the reddening displayed in the innermost bins of the $u - b_J$ colours correspond to an increase to the blue in the $b_J - r$ colours, further suggesting it is a small-scale measurement artefact rather than due to dust. In the lower panel of Fig. 5.11, we translate the bootstrapped error bars into limits on absorption by dust around galaxies. The solid $u - b_J$ and dashed $b_J - r$ lines in the lower panel represent the $1\sigma$ limit on the anti-correlation due to dust allowed by our colour limits. The anti-correlation between QSOs and galaxies measured in Fig. 5.8 is plotted for comparison. The $1\sigma$ limits allowed by dust are insufficent to account for the anti-correlation between galaxies and QSOs, in fact, the $b_J - r$ limits are ruled out at a $2\sigma$ level and could only account for about 30 per cent of the anti-correlation between QSOs and galaxies (out to 10 arcminutes). Acknowledging the proviso that we have only considered dust laws representative of our own Galaxy, we conclude that dust around galaxies can not account for the anti-correlation between QSOs and galaxies but can we find definitive evidence that gravitational lensing is responsible?

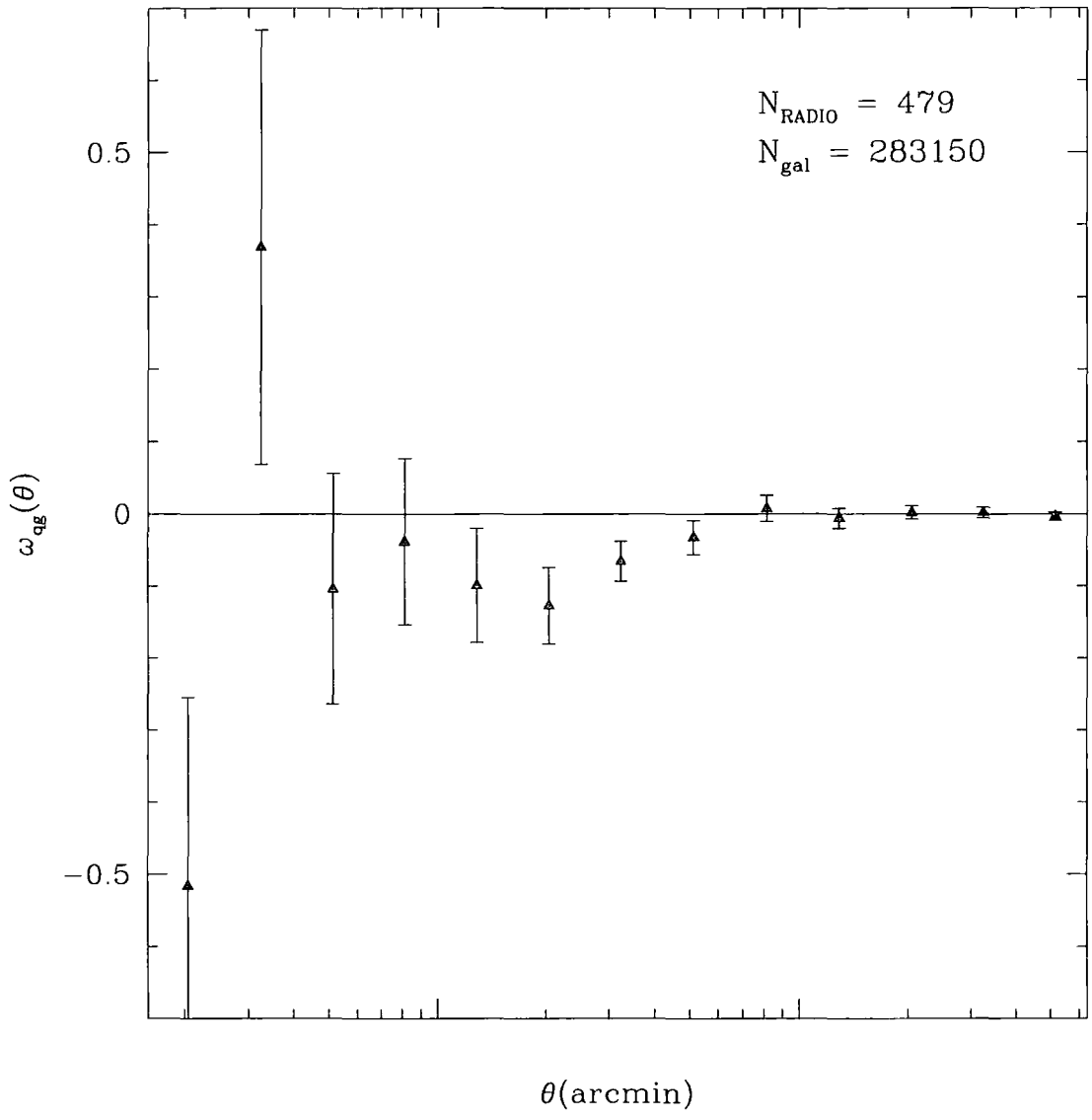


Figure 5.12: This figure shows the cross-correlation between galaxies and a radio-selected subsample of the 2QZ. All cross-correlations are carried out against SDSS EDR galaxies (in the region of the NGC 2QZ strip) or against APM galaxies (in the region of the SGC 2QZ strip). The points mark the cross-correlation signal for radio-loud 2QZ QSOs against galaxies (for the NGC and SGC strips combined). Error bars represent $1\sigma$ jackknife errors. Labels note the number of galaxies and Radio-loud QSOs present within the confines of the 2QZ boundaries.

#### 5.4.4.2 Double Magnification Bias Revisited

In Section 5.3 of this chapter, we introduced the concept of double magnification bias (Borgeest, von Linde & Refsdal 1991) and presented arguments that at the simplest level we would not expect this signature of lensing to be seen in a radio-loud subsample of the 2QZ but that some authors (Wyithe et al. 2003) have argued that it might arise in subsamples selected in passbands that correlate but with a great deal of scatter, which we showed to be the situation with radio-loud optically-identified 2QZ QSOs. In Fig. 5.12 we show the cross-correlation of our galaxy samples with radio-loud QSOs. We measure a strong anti-correlation between these samples out to 5 arcminute scales; $\omega_{qg}(< 5') = -0.08$. This anti-correlation is significant at the $3.3\sigma$ level (99.9 per cent significance). It may be the case that this strong anti-correlation (over and above that seen in the full 2QZ optically-selected QSO sample) merely arises as the effect of a particularly shallow slope in the radio-loud QSO number-counts, rather than any genuine double magnification bias, although Fig. 5.4 would suggest not. Tentatively, we suggest that double magnification bias is a mechanism that could induce a strong signal in a subsample of QSOs selected in both the radio and optical. We have no other explanation, beyond a statistical fluke, for the strong dearth of radio-loud QSOs around foreground galaxies.

#### 5.4.4.3 Evolution of the Cross-Correlation Signal With Magnitude and Redshift

Perhaps the key prediction of the statistical lensing hypothesis as regards the relative distribution of QSOs around galaxies is the effect of the magnitude of the QSO sample on the cross-correlation signal. Equation 5.2 predicts an enhancement of QSOs near tracers of foreground lenses when the slope of the QSO number-magnitude counts is greater than 0.4 and a deficit of QSOs around the same lenses when the slope is less than 0.4. In Chapter 2, we displayed the QSO number-counts by magnitude. The "knee" of the magnitude distribution, where the slope is 0.4, lies around $b_J = 19.1$ to $b_J = 19.6$. A very simple model would suggest that, under the assumption that lensing magnifies QSOs about a magnitude fainter than the sample limit, we might expect a positive correlation between QSOs and galaxies up to a (QSO) $b_J$ magnitude of around 18.1-18.6, no correlation between QSOs and galaxies in the range 18.1-18.6 to the knee of the magnitude counts, and an anti-correlation between QSOs and galaxies fainter than about $b_J = 19.6$.

At the time of writing, no author has yet shown the evolution of the QSO-galaxy cross-

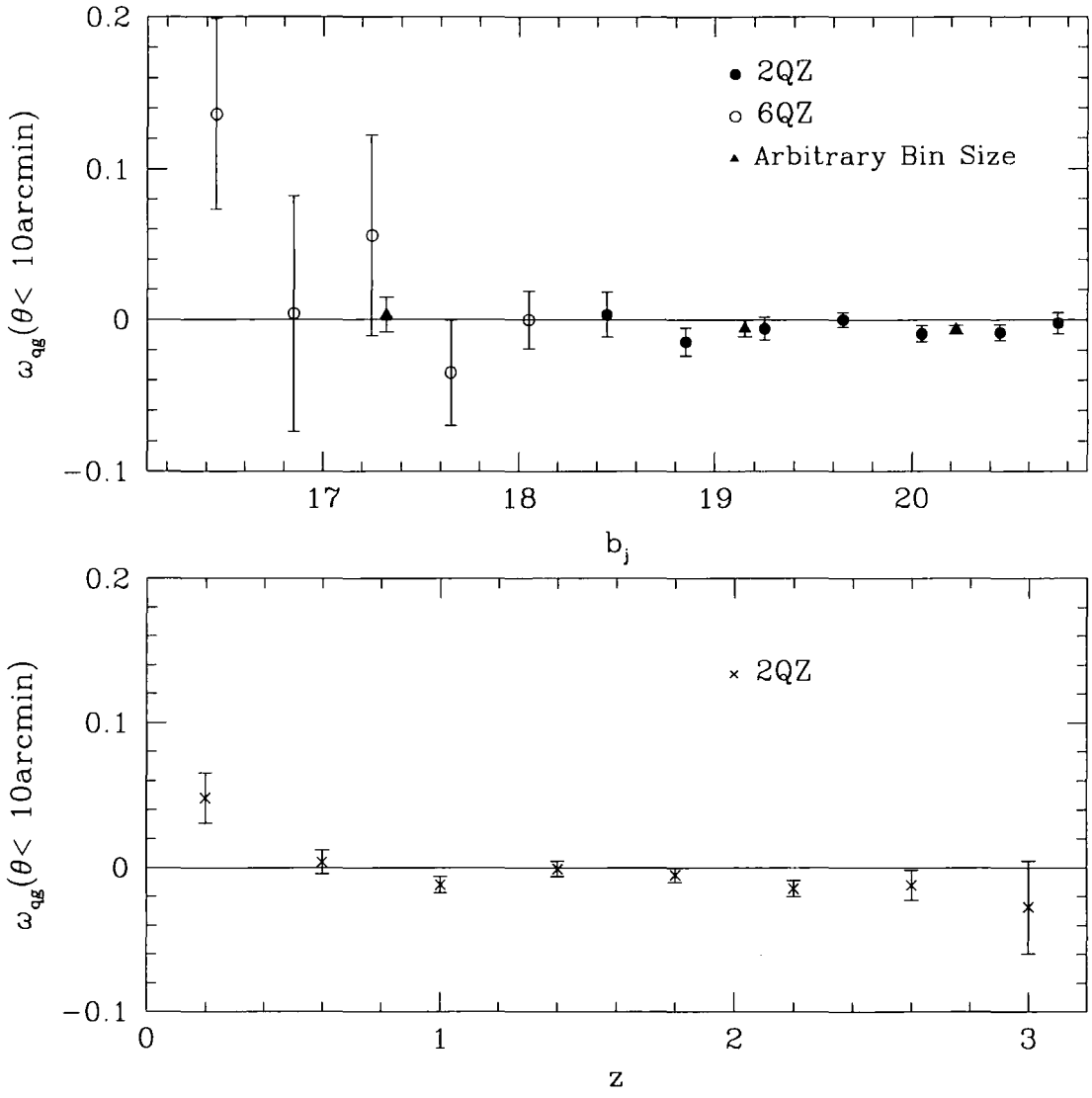


Figure 5.13: This figure concerns how the cross-correlation of 2QZ QSOs against SDSS EDR galaxies (in the region of the NGC 2QZ strip) and against APM galaxies (in the region of the SGC 2QZ strip) depends on $b_J$ magnitude and redshift. The upper panel displays the cross-correlation signal measured out to 10 arcminutes for subsamples of 2QZ QSOs in bins of extent $b_J = 0.4$mag. 2QZ QSOs are marked as filled circles, 6QZ as open circles. We have also divided the data into three bins of "arbitrary" size, representing the bright-end, the knee and the faint-end of the QSO number-magnitude counts. These points are marked as triangles. The lower panel displays the cross-correlation signal measured out to 10 arcminutes for subsamples of 2QZ QSOs in bins of redshift extent $z = 0.4$. Error bars represent $1\sigma$ jackknife errors.

correlation signal with magnitude from a positive correlation at bright QSO magnitudes to an anti-correlation at faint magnitudes, although some authors have shown the transition from a positive correlation to zero correlation (Williams & Irwin 1998, Gaztañaga 2003). This is mainly a problem of QSO sampling - as yet no single survey spans the QSO magnitude distribution in a manner that produces significant numbers of QSOs at both bright and faint magnitudes. The combined area and depth of the 2QZ and 6QZ allows us to trace the evolution of the cross-correlation between QSOs and galaxies with magnitude for the first time. In the upper panel of Fig. 5.13 we show the evolution of the QSO-galaxy cross-correlation signal (measured out to 10 arcminutes) in (differential) 0.4 mag bins spanning the range $16.25 < b_J < 20.65$. QSOs are taken from the 2QZ for $b_J > 18.25$ and from the 6QZ survey for $b_J < 18.25$. The system of binning was chosen so that significant numbers of QSOs would still be in bright bins, while ensuring that no single bin spanned the two surveys. This leaves two smaller bins with data in them spanning $20.65 < b_J < 20.85$ and $16 < b_J < 16.25$, completing the range of the surveys. The brightest bin has essentially no data in it and is not plotted in Fig. 5.13. There is a loose trend in the data suggesting that the brightest QSOs are positively correlated with our galaxy samples (at the $2.2\sigma$ level for $b_J < 16.65$) and the faintest QSOs are anti-correlated with our galaxy samples (at the $2.3\sigma$ level for $b_J > 19.85$) and there is no significant result for the remainder of the magnitude range. We do not trust the signal in the brightest bin - there are only a handful of objects in this bin and the positive correlation is likely a statistical fluke. We have additionally created three "arbitrary" bins, chosen both to increase the number of QSOs in the brighter bins and to reflect the "bright" QSO sample, the "knee" QSO sample and the "faint" QSO sample, or where statistical lensing would predict a positive correlation, no correlation and an anti-correlation. These "arbitrary" bins are plotted as triangles in Fig. 5.13. Although there remains a significant anti-correlation (-0.006 at $2.1\sigma$) in the faintest bin ($b_J > 19.60$) and the signal has an insignificant trend towards the positive at brighter magnitudes, we cannot say with any confidence that we have detected the full evolution of the cross-correlation signal with magnitude predicted by statistical lensing. This is not necessarily surprising, as there are very few objects in the 6QZ and only 256 QSOs brighter than ($b_J = 18.1$) that meet our usual redshift and identification quality selection criteria. Nevertheless, we do see a trend away from anti-correlations for brighter samples, in line with the statistical lensing hypothesis.

In the lower panel of Fig. 5.13, we show the evolution of the cross-correlation of 2QZ QSOs against our combined galaxy samples with redshift in differential bins of $z = 0.4$.

The simple lensing model outlined in Section 3.4 predicts no detectable trend in the cross-correlation signal with QSO redshift. We might expect a stronger signal at larger redshifts, simply because we will on average sample fainter QSOs in these bins and will thus observe the signature of changes in magnitude in the redshift distribution. Indeed, we see a reasonably consistent anti-correlation for all redshift bins and the signal is slightly stronger at high redshifts but not significantly so. The lowest redshift QSOs are significantly correlated with our galaxy samples, no doubt due to genuine clustering at these redshifts, justifying our cut of $z < 0.4$ in other analyses throughout this chapter.

In this section, we have shown that there is a significant anti-correlation between faint QSOs and galaxies. This signal does not arise as a systematic effect of how the QSO samples are selected. We rule out the possibility that the majority of the signal is due to dust distributed around galaxy groups at the $2\sigma$ level by comparing $b_J - r$ QSO colours in the field and close to galaxies, with the caveat that the dust is distributed randomly and obeys dust laws typical for our Galaxy. We have shown that the anti-correlation is consistent with the predictions of statistical lensing, both in terms of the evolution of the correlation signal with QSO magnitude and the effect of double magnification bias on the radio-loud 2QZ QSO sample. We will now model the anti-correlation on the basis that it is due to statistical lensing and consider cosmological implications.

## 5.5 Statistical Lensing Models

In this section, we outline two lensing models we shall use in describing the anti-correlation between QSOs and galaxies. Both models work by comparing the cross-correlation of QSOs and galaxies with the auto-correlation of galaxies to determine the biasing of galaxies with respect to the mass that the QSO light traces. The first, based mainly on the work of Williams & Irwin (1998) uses a simple linear biasing prescription to relate fluctuations in the foreground mass distribution to QSO magnification. Williams & Irwin (1998) have shown that this simple model agrees well with more complicated simulations (Dolag & Bartelmann 1997, Sanz, Martínez-González & Benítez 1997). The second model, due to Gaztañaga (2003), is appropriate for both linear and non-linear biasing prescriptions and makes excellent use of the formalism that has evolved over the last few decades to model galaxy-galaxy auto-correlations - see, for instance, (Peebles & Groth 1976, Groth & Peebles 1977, Fall 1979, Fry 1986, Balian & Schaeffer 1989, Gaztañaga & Yokoyama 1993, Gaztañaga 1994) and many other authors referenced therein.

### 5.5.1 Linear Bias Model

The convergence of lensing matter is defined as (see Appendix C)

$$\kappa(\theta) = \frac{\Sigma(D_l, \theta)}{\Sigma_{cr}(D_l, D_s)} \tag{5.8}$$

where $\Sigma(D_l, \theta)$ is the surface mass density of the lensing material. The critical mass surface density is defined

$$\Sigma_{cr}(D_l, D_s) = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}} \tag{5.9}$$

for (angular diameter) lens distance $D_l$, source distance $D_s$ and lens-source separation $D_{ls}$, where $c$ is the speed of light and $G$ is the gravitational constant. Angular diameter distances can be calculated from redshift (see Appendix A). Hence $\Sigma_{cr}(z, z_s)$ is a function of the redshift of background source QSOs ($z_s$) and of the foreground lensing matter ($z$). We take $z_s = 1.5$, the median redshift of the 2QZ, for the background source redshift. Although this seems a rather extreme approximation of the actual distribution of QSOs, Bartelmann & Schneider (2001) suggest it is fair. It is not completely straightforward to calculate the angular diameter distance between two distinct redshifts. We determine the term $D_{ls}$ using the approach of Kayser, Helbig & Schramm (1997). We calculate all distances in this thesis assuming a smooth (homogenous) universe.

Williams & Irwin (1998) model the dark matter as a uniform, smooth-universe slab that extends over the redshift range $z = 0$ to $z = z_{max}$. Lensing arises in the slab due to fluctuations in the matter density around the uniform mean density of the slab. We will characterise the mean projected fluctuations from $z = z_{max}$ to $z = 0$ within the slab by $\delta(\theta)$ (the projected angular density contrast of the lensing matter). The effective convergence *due to angular fluctuations in the lensing matter* is then $\kappa_{eff}(\theta) = \bar{\kappa}\delta(\theta)$, where $\bar{\kappa}$ is the mean convergence of the lensing slab. We define $\bar{\kappa}$ using Equation 5.9 as $\bar{\kappa} = \bar{\Sigma}(z)/\Sigma_{cr}(z, z_s)$, where $z$ is now the redshift of the lens at angular diameter distance $D_l$.

By definition, the average mass density of the Universe at redshift $z = 0$ is $\rho_{crit}\Omega_m$. Note that $\rho_{crit}$ is not related to $\Sigma_{cr}(z, z_s)$, the former being the critical density of the Universe, the latter being the critical density necessary for strong lensing to occur (see Appendix A and Appendix C). In the matter-dominated era, the density of the Universe is inversely proportional to the cube of the scale factor (see Appendix A), so the average mass density of the Universe at redshift $z$ is $\rho_{crit}\Omega_m(1+z)^3$. The extent of a lensing slab

is the time taken by a light ray to traverse it, $cdt$, where $t(z)$ is a function of redshift, so the mass density of an infinitely thin slab of matter at redshift $z$ is

$$\bar{\Sigma}(z) = \rho_{crit}\Omega_m(1+z)^3 cdt = \frac{3H_0^2 c}{8\pi G}\Omega_m(1+z)^3 dt \tag{5.10}$$

and, combining Equation 5.10 with equations from the previous two paragraphs, the effective convergence due to angular fluctuations in the mass density of a slab of matter that extends from the observer at $z = 0$ to redshift $z_{max}$ is

$$\kappa_{eff}(\theta) = \bar{\kappa}\delta(\theta) = \frac{\bar{\Sigma}(z)}{\Sigma_{cr}(z, z_s)}\delta(\theta) \tag{5.11}$$

where, we remember, $\delta(\theta)$ is the projected angular density contrast of the lensing matter, and

$$\bar{\kappa} = \frac{3H_0^2 c}{8\pi G}\Omega_m \int_0^{z_{max}} \frac{(1+z)^3 \frac{dt}{dz} dz}{\Sigma_{cr}(z, z_s)} \tag{5.12}$$

The factor $dt/dz$ is derived in Appendix A.

Equation 5.11 allows us to calculate the convergence due to angular density fluctuations in the lensing slab in terms of the properties of the lensing slab ($z_{max}$, $\delta$) the lensed source ($z_s$) and cosmological parameters ($\Omega_m, \Omega_\Lambda, H_0$). We now relate this convergence to the correlation functions.

When we centre on a galaxy and count the average number of galaxies around it, the enhancement of the galaxy number density over the average can be expressed as (Bartelmann 1995)

$$\frac{n_G(\theta) - \langle n_G(\theta)\rangle}{\langle n_G(\theta)\rangle} = b\delta(\theta) \tag{5.13}$$

where $b$ is the (scale-independent) linear bias parameter. We will call this $\omega_{1G}(\theta)$, as it is not the same thing as the galaxy-galaxy auto-correlation but approximates it in the case of searching around a single galaxy. The projected angular density contrast of the lensing matter, $\delta(\theta)$, from Equation 5.13 can be substituted into Equation 5.11 to give

$$\kappa_{eff}(\theta) = \bar{\kappa}\delta(\theta) = \frac{\bar{\kappa}\omega_{1G}(\theta)}{b} \tag{5.14}$$

where $\bar{\kappa}$, the mean convergence of the lensing slab, is given by Equation 5.12.

The last step in this modelling process is to express the convergence due to the projected density fluctuations within the lensing slab in terms of the magnification of the

lensed sources. The magnification may be written in terms of the determinant of the lensing matrix (see Appendix C). For point source images (which exhibit no shear) there is a direct relation between magnification and convergence

$$\mu = \frac{1}{(1-\kappa)^2} = 1 + 2\kappa + 3\kappa^2 + 4\kappa^3 + ..... \tag{5.15}$$

In the weak lensing approximation, the convergence is small and the second order terms are negligible. So, the magnification due to lensing by fluctuations in the matter distribution of a slab extending out from $z = 0$ to $z = z_{max}$ is, substituting for $\kappa_{eff}$ from Equation 5.14

$$\mu(\theta) = 1 + 2\kappa_{eff}(\theta) = 1 + 2\frac{\bar{\kappa}\omega_{1G}(\theta)}{b} \tag{5.16}$$

The enhancement of the QSO number density over the average around any galaxy is given, for the QSO integrated number-counts with slope $\alpha$, by Equation 5.2, above. So, in the weak lensing approximation, we have

$$1 + \omega_{1Q}(\theta) = q = \mu^{2.5\alpha-1} = \left(1 + \frac{2\bar{\kappa}\omega_{1G}(\theta)}{b}\right)^{2.5\alpha-1} \tag{5.17}$$

Again noting that the convergence is small in the weak lensing approximation and taking the linear terms of a Taylor expansion

$$1 + \omega_{1Q}(\theta) = 1 + (2.5\alpha - 1)\frac{2\bar{\kappa}\omega_{1G}(\theta)}{b} \tag{5.18}$$

Although this is an approximation based on counting the distribution of galaxies and QSOs around a particular galaxy, if we ensure that the same galaxies are being used when calculating the correlation function it is fair to write

$$\omega_{gq}(\theta) = (2.5\alpha - 1)\frac{2\bar{\kappa}\omega_{gg}(\theta)}{b} \tag{5.19}$$

where $\omega_{gq}(\theta)$ is now the cross-correlation of QSOs and galaxies and $\omega_{gg}(\theta)$ is the galaxy-galaxy auto-correlation. Other than the bias parameter, the cosmology is contained in the (dimensionless) mean convergence $\bar{\kappa}$ defined in Equation 5.12. The QSO number-counts are included as the power-law slope, $\alpha$. As mentioned at the start of this section, this model compares extremely well with detailed simulations (Williams & Irwin 1998). It

is relatively easy to extend Equation 5.19 to a simple model of scale-dependent bias, by expressing $\omega_{gq}(\theta)$ and $\omega_{gg}(\theta)$ as power laws with angular dependency; $\omega(\theta) = A\theta^{\gamma}$; and extrapolating the angular dependence, if any, of the bias.

### 5.5.2 General Model

Gaztañaga (2003) has produced a comprehensive model of galaxy-QSO cross-correlations based on standard, widely accepted galaxy clustering theory. The model should also be valid for a non-linear biasing prescription, which is important in this work as bias is likely scale-dependent on small ($< 10\,h^{-1}$Mpc) scales (Blanton et al. 1999) and N-body simulations suggest the matter distribution becomes highly non-linear on sufficiently small scales (Hamilton et al. 1991, Peacock & Dodds 1996, Smith et al. 2003). The model also allows for realistic redshift distributions for the QSO and galaxy distributions, meaning there is no ambiguity of just how far in redshift to integrate the lensing matter out to. In this section, we reproduce the theory of Gaztañaga (2003) from a number of different sources to illuminate its origins and relevance.

#### 5.5.2.1 Mass fluctuations

Peebles & Groth (1976), (see also page 216 of Peebles 1980, or Equation 7.72 of Peebles 1993) have shown that for (three-dimensional) correlation functions that are well-fit by the power-law form

$$\xi(r) = \left(\frac{r_0}{r}\right)^{\gamma} \tag{5.20}$$

(where $r_0$ is the scale-length of the fit), the volume-averaged integral over a sphere of radius R (with top-hat smoothing) is

$$\bar{\xi}(R) = \frac{1}{V^2}\int_0^R dV_1 dV_2 \xi(|r_1 - r_2|) = \left(\frac{r_0}{R}\right)^{\gamma} \frac{72}{(3-\gamma)(4-\gamma)(6-\gamma)2^{\gamma}} \tag{5.21}$$

Often, for comparison to simulations, mass fluctuations are measured on a scale corresponding to a sphere of $8\,h^{-1}$Mpc and characterised by the parameter $\sigma_8$ (see, e.g.; White, Efstathiou & Frenk 1993; and references therein). We, however, will measure on a scale of $0.1\,h^{-1}$Mpc (consider Gaztañaga 1995)

$$\bar{\xi}(R) = \sigma_{0.1}^2 \left(\frac{0.1\,h^{-1}\mathrm{Mpc}}{R}\right)^{\gamma} \tag{5.22}$$

with

$$\sigma_{0.1}^2 = \frac{72}{(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma}\left(\frac{r_0}{0.1\,h^{-1}\mathrm{Mpc}}\right)^\gamma \tag{5.23}$$

as at the median depth of our galaxy samples ($z \sim 0.15$), 1 arcminute is about $0.1\,h^{-1}$Mpc. These equations express the three-dimensional correlations of mass (or, for our purposes, underlying lensing matter) as a power law, and all of the terms, $r_0, \gamma, \sigma_{0.1}$ etc., refer to the underlying lensing mass. Gaztañaga (2003), (see the appendix of Gaztañaga 2003 and also Fry & Gaztañaga 1993), has shown that a good power-law expression for both linear and non-linear biasing of galaxies is

$$b(R) = b_{0.1}\left(\frac{0.1\,h^{-1}\mathrm{Mpc}}{R}\right)^{\gamma_b} \tag{5.24}$$

and further shows that the galaxy-averaged correlation function can thus be expressed in terms of the mass-averaged function of Equation 5.22 as

$$\bar{\xi}_G(R) = b_{0.1}^2\sigma_{0.1}^2\left(\frac{0.1\,h^{-1}\mathrm{Mpc}}{R}\right)^{\gamma+2\gamma_b} \tag{5.25}$$

#### 5.5.2.2 Projection

Limber (1953) first discussed the projection of three-dimensional density functions into purely angular distributions and his theory was developed for cosmological applications by many other authors (Totsuji & Kihara 1969, Phillipps et al. 1978, Fall 1979). Groth & Peebles (1977) proposed the "$\epsilon$" form of the correlation function to describe the evolution of clustering with redshift

$$\xi(r) = \left(\frac{r_0}{r}\right)^\gamma (1+z)^{-3+\epsilon} \tag{5.26}$$

If the universe is composed of physically invariant clusters then $\epsilon = 0$. This is often called "stable clustering" as the cluster evolution is fixed in proper coordinates. Note that the clustering appears to grow with time for $\epsilon = 0$, as the expansion of the universe leads to an increase in the size of the voids between the physically invariant clusters. If $\epsilon = \gamma - 3$, the clustering is fixed in comoving coordinates and clusters expand in check with the universe, meaning clustering does not increase with time. If $\epsilon > 0$ then clustering

grows with time in proper coordinates, as would be expected due to gravitational infall. N-body simulations suggest that, for a ΛCDM cosmology, there is no single value of $\epsilon$ that fits the evolution of the clustering of dark matter out to $z \sim 6$, although $\epsilon \sim 0.8$ is the best approximation. At smaller scales, $0 \lesssim z \lesssim 0.5$, linear theory predicts $\epsilon = -0.15$ for an $\Omega_m = 0.2, \Omega_\Lambda = 0.8$ cosmology (see table 2 of Carlberg et al. 2000). The observational evidence does not constrain $\epsilon$ particularly well (Wilson 2003), mainly due to difficulties in selecting uniform samples and in converting measurements of the galaxy-galaxy auto-correlation into the correlation of the underlying mass. It has been suggested, on theoretical considerations, that stable clustering should hold on small scales, in the highly non-linear regime (Hamilton et al. 1991, Peacock & Dodds 1994), a proposition that may or may not be true (Smith et al. 2003). For the purposes of this model, we shall assume that $\epsilon = 0$ for small-scale correlations of lensing matter, at least at the low redshifts ($z \sim 0.15$), where our galaxy samples reside.

For our purposes, Equation 5.26, as was the case for Equation 5.22, represents the three-dimensional correlation function of lensing dark matter. The two-dimensional projection of Equation 5.26 (i.e. the angular correlation function) is (Groth & Peebles 1977, also section 56 of Peebles 1980)

$$\omega(\theta) = A\theta^{1-\gamma} \tag{5.27}$$

where

$$A = {r_0}^\gamma \frac{\Gamma(1/2)\Gamma(\gamma/2-1/2)}{\Gamma(\gamma/2)} \frac{\int_0^\infty \chi^{5-\gamma} d\chi \phi(\chi)^2 (1+z)^{\epsilon-3}}{[\int_0^\infty \chi^2 d\chi \phi(\chi)]^2} \tag{5.28}$$

for a flat cosmology, where $\chi$ is the comoving distance (see Appendix A) and $\Gamma$ is the gamma function. The comoving distance is easily converted into redshift for a given cosmology (see Appendix A). When measuring correlation functions we express $\theta$ in arcminutes, so it is worth noting that $\theta$ is expressed in radians in Equation 5.27.

In terms of survey redshifts, the selection function is $\int_0^\infty \chi^2 d\chi \phi(\chi) = \int_0^\infty \frac{dN}{dz} dz$ (see, e.g., Efstathiou et al. 1991 or Wilson 2003), where $N$ is the number of objects in the survey. Substituting this into Equation 5.28 and transforming into comoving coordinates

$$A = {r_0}^\gamma \frac{\Gamma(1/2)\Gamma(\gamma/2-1/2)}{\Gamma(\gamma/2)} \frac{\int_0^\infty \chi^{1-\gamma} \left(\frac{dN}{d\chi}\right)^2 d\chi (1+z)^{\epsilon-3}}{[\int_0^\infty \frac{dN}{d\chi} d\chi]^2} \tag{5.29}$$

We normalise the denominator so that $\int_0^\infty \frac{dN}{d\chi} d\chi = 1$. We can determine the comoving selection function (also called the *weight function*)from the survey redshift distributions (plotted in Fig. 2.6 of Chapter 2) given that

$$\int_0^{\chi'} \frac{dN}{d\chi} d\chi = \int_0^{z'} \frac{dN}{dz} dz \tag{5.30}$$

Finally, if we substitute into Equation 5.29 for $r_0$ from Equation 5.23 we find the angular mass correlation is

$$\omega = \sigma_{0.1}^2 A(0.1\,h^{-1}\mathrm{Mpc}, \gamma)\theta^{1-\gamma} \tag{5.31}$$

where

$$A(0.1\,h^{-1}\mathrm{Mpc}, \gamma) = B_{0.1}(\gamma) \int_0^\infty d\chi \left(\frac{dN}{d\chi}\right)^2 \chi^{1-\gamma}(1+z)^{\epsilon-3} \tag{5.32}$$

and

$$B_{0.1}(\gamma) = \frac{(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma}{10^\gamma 72} \frac{\Gamma(1/2)\Gamma(\gamma/2-1/2)}{\Gamma(\gamma/2)} \tag{5.33}$$

The factor $10^\gamma$ in the denominator arises as $r_0$ is expressed in $h^{-1}$Mpc in Equation 5.29 but $r_0$ is expressed in $0.1\,h^{-1}$Mpc in Equation 5.23. We will also consider a model where this factor is $5^\gamma$ and the bias and mass fluctuations are considered to be measured on $0.2\,h^{-1}$Mpc scales, to compare our results to Gaztañaga (2003). Even though Gaztañaga (2003) considered a galaxy distribution with the same scale as ours, they also considered a distribution with a median redshift of 0.35, which would have an angular extent $0.3\,h^{-1}$Mpc at 1 arcminute, so chose a scale of $0.2\,h^{-1}$Mpc for their analysis. As this factor in the denominator makes $B(\gamma)$ dependent on the scale of the mass fluctuations that are being considered, we will denote it $B_{0.1}(\gamma)$, to remind the reader that it is scale-dependent. We will denote $A(\gamma)$ by $A(0.1\,h^{-1}\mathrm{Mpc}, \gamma)$ for the same reason.

#### 5.5.2.3 Magnification Bias

It is simple to express the projected angular fluctuation (or density contrast, $\delta$) of *galaxies* integrated along any line of sight, in terms of the density contrast at any point along that line of sight and the number of galaxies, as a function of comoving coordinates

$$\delta_G(\theta) = \int_0^{\chi_H} d\chi \frac{dN_G}{d\chi} \delta_G(\theta, \chi) \tag{5.34}$$

where we have specified $\frac{dN}{d\chi}$ with a subscript "G" as the galaxy selection function in $\chi$-space and the subscript "H" refers to the cosmological horizon.

Projected fluctuations in QSO numbers due to lensing (discussed in length in this thesis as magnification bias) are characterised by the enhancement factor, $q$, of Equation 5.2, a measure of the number of lensed QSOs compared to the unlensed population

$$\delta_\mu(\theta) = \frac{n_Q(\theta) - \langle n_Q(\theta) \rangle}{\langle n_Q(\theta) \rangle} = q - 1 = \mu^{2.5\alpha - 1} - 1 \tag{5.35}$$

where the subscript $\mu$ denotes the density contrast due to magnification bias and $\alpha$ is the power-law slope of the QSO number-magnitude counts. In the weak lensing approximation (see Equation 5.15) we can relate the fluctuations due to lensing to the convergence (Bartelmann 1995, Bartelmann & Schneider 2001, Menard & Bartelmann 2002)

$$\delta_\mu(\theta) = (2.5\alpha - 1)\delta\mu = 2(2.5\alpha - 1)\kappa_{eff}(\theta) \tag{5.36}$$

where $\kappa_{eff}(\theta)$, the convergence due to fluctuations in the expanse of lensing matter, is determined from the cosmological form of Poisson's Equation, as discussed in Appendix C (Bernardeau, van Waerbeke & Mellier 1997, Bartelmann & Schneider 2001, Guimãraes 2002, Menard & Bartelmann 2002, Gaztañaga 2003)

$$\kappa_{eff}(\theta) = \int_0^{\chi_H} d\chi \frac{dN_Q}{d\chi} \kappa_{eff}(\theta, \chi) \tag{5.37}$$

$$= \frac{3{H_0}^2 \Omega_m}{2c^2} \int_0^{\chi_H} d\chi (1+z)\chi \delta(\theta, \chi) \int_\chi^{\chi_H} d\chi' \frac{\chi' - \chi}{\chi'} \frac{dN_Q}{d\chi'} \tag{5.38}$$

where $\frac{dN_Q}{d\chi}$ is the QSO selection function in comoving coordinates, determined as for Equation 5.30. The term $\delta(\theta, \chi)$, the density contrast of the underlying lensing matter, enters Equation 5.38 through Poisson's Equation (see Appendix C). The density contrast due to magnification bias may then be expressed as

$$\delta_\mu(\theta) = (2.5\alpha - 1) \frac{3{H_0}^2 \Omega_m}{c^2} \int_0^{\chi_H} d\chi (1+z)\chi \delta(\theta, \chi) \int_\chi^{\chi_H} d\chi' \frac{\chi' - \chi}{\chi'} \frac{dN_Q}{d\chi'} \tag{5.39}$$

$$= \int_0^{\chi_H} d\chi \varepsilon(\chi) \delta(\theta, \chi) \tag{5.40}$$

Where we have deliberately expressed the lensing density contrast in a form similar to Equation 5.34, allowing us to define the lensing efficiency (Bernardeau, van Waerbeke & Mellier 1997) as

$$\varepsilon(\chi) = (2.5\alpha - 1)\frac{3{H_0}^2\Omega_m}{c^2}(1+z)\chi\int_\chi^\infty d\chi'\frac{\chi'-\chi}{\chi'}\frac{dN_Q}{d\chi'} \tag{5.41}$$

By analogy with Equation 5.34, the $\varepsilon(\chi)$ is the selection function of the lensing effect in comoving coordinates.

#### 5.5.2.4 Correlation Functions

As outlined in Gaztañaga (2003), the equations derived in this section allow us to model the galaxy-galaxy auto-correlation and galaxy-QSO cross-correlation. Equations 5.25,5.31 and 5.32 dictate that

$$\omega_{gg}(\theta) = \sigma_{0.1}^2 b_{0.1}^2 \theta^{1-\gamma-2\gamma_b} A_{gg}(0.1\,h^{-1}\mathrm{Mpc}) \tag{5.42}$$

$$A_{gg}(0.1\,h^{-1}\mathrm{Mpc}) = B_{0.1}(\gamma+2\gamma_b)\int_0^\infty d\chi\left(\frac{dN_G}{d\chi}\right)^2 \chi^{1-\gamma-2\gamma_b}(1+z)^{\epsilon-3} \tag{5.43}$$

where $B_{0.1}(\gamma)$ is given by Equation 5.33. The QSO-galaxy cross-correlation is similarly characterised by Equations 5.25, 5.31, 5.32 and 5.41 as

$$\omega_{gq}(\theta) = \sigma_{0.1}^2 b_{0.1} \theta^{1-\gamma-\gamma_b} A_{gq}(0.1\,h^{-1}\mathrm{Mpc}) \tag{5.44}$$

$$A_{gq}(0.1\,h^{-1}\mathrm{Mpc}) = B_{0.1}(\gamma+\gamma_b)\int_0^\infty d\chi\left(\frac{dN_G}{d\chi}\right)\varepsilon(\chi)\chi^{1-\gamma-\gamma_b}(1+z)^{\epsilon-3} \tag{5.45}$$

where $\varepsilon(\chi)$, the lensing selection function, is expressed in Equation 5.41. $\varepsilon(\chi)$ is typically less than 1 per cent of $\frac{dN_G}{d\chi}$, which has often led to the prediction that $\omega_{gq}$ should only be observed to be a tiny fraction of $\omega_{gg}$ (Dolag & Bartelmann 1997).

The measurables are $\alpha$ (the slope of the QSO number-magnitude counts) and $\gamma$ and $\gamma_b$ which are easily determined from the slopes of the correlation functions. If $\gamma_{gg}$ and $\gamma_{gq}$ are the slopes of the galaxy-galaxy auto-correlation and QSO-galaxy cross-correlation respectively, then, using Equations 5.43 and 5.45

$$1+\gamma_{gg}-2\gamma_{gq} = 1+(1-\gamma-2\gamma_b-2)+2(1-\gamma-\gamma_b) = \gamma \tag{5.46}$$

$$\gamma_{gq}-\gamma_{gg} = 1-\gamma-\gamma_b-1+\gamma+2\gamma_b = \gamma_b \tag{5.47}$$

Hence, the bias parameter and $\sigma_{0.1}$ can be determined from

$$b_{0.1} = \frac{\omega_{gg}(\theta)}{\omega_{gq}(\theta)} \frac{A_{gq}(0.1\,h^{-1}\mathrm{Mpc})}{A_{gg}(0.1\,h^{-1}\mathrm{Mpc})} \theta^{\gamma_b} \quad (5.48)$$

$$\sigma^2_{0.1} = \frac{\omega_{gq}(\theta)^2}{\omega_{gg}(\theta)} \frac{A_{gg}(0.1\,h^{-1}\mathrm{Mpc})}{A^2_{gq}(0.1\,h^{-1}\mathrm{Mpc})} \theta^{\gamma-1} \quad (5.49)$$

An advantage of this modelling approach over the simpler approach of (Williams & Irwin 1998), discussed in Section 5.5.1, is that the estimation of the cosmological parameters depends both on the slope of the correlation functions, as well as their amplitude. The slope might be well-constrained even when the amplitude of the cross-correlation signal at a given point is in doubt. The two approaches (Williams & Irwin 1998, Gaztañaga 2003) are thus also complementary, using different aspects of the correlation measurements, as well as different models of the lensing matter, to predict cosmological parameters.

## 5.6 Results

In Fig. 5.14 we plot the auto-correlation of galaxies combined from the SDSS EDR in the NGC 2QZ strip and the APM Survey in the SGC 2QZ strip. Fig. 5.14 also shows the anti-correlation between QSOs and galaxies discussed throughout this chapter. We have plotted the result for the QSO sample fainter than $b_J = 19.6$, to ensure we are sampling QSOs from a region of the number-magnitude counts fainter than the "knee" of the distribution (see Section 5.4.4.3). We fit simple power-laws to the two correlation functions, finding that

$$\omega_{gq}(\theta) = -0.024 \pm^{0.008}_{0.007} \theta^{-1.0\pm0.3} \quad (5.50)$$

$$\omega_{gg}(\theta) = 0.330 \pm^{0.015}_{0.014} \theta^{-0.76\pm0.04} \quad (5.51)$$

where $\theta$ is expressed in arcminutes, 1 arcminute being about $0.1\,h^{-1}$Mpc at the median redshift of our galaxy data. In both cases, a simple $\chi^2$ fit suggests excellent power-law fits to the data, with a reduced $\chi^2$ value of 1.6 for the galaxy-galaxy correlation fit and 1.4 for the QSO-galaxy anti-correlation fit. The result for the slope of the galaxy-galaxy auto-correlation is in excellent agreement with recent small-scale measurements for galaxies with similar limiting flux to the sample used here (Connolly et al. 2002).

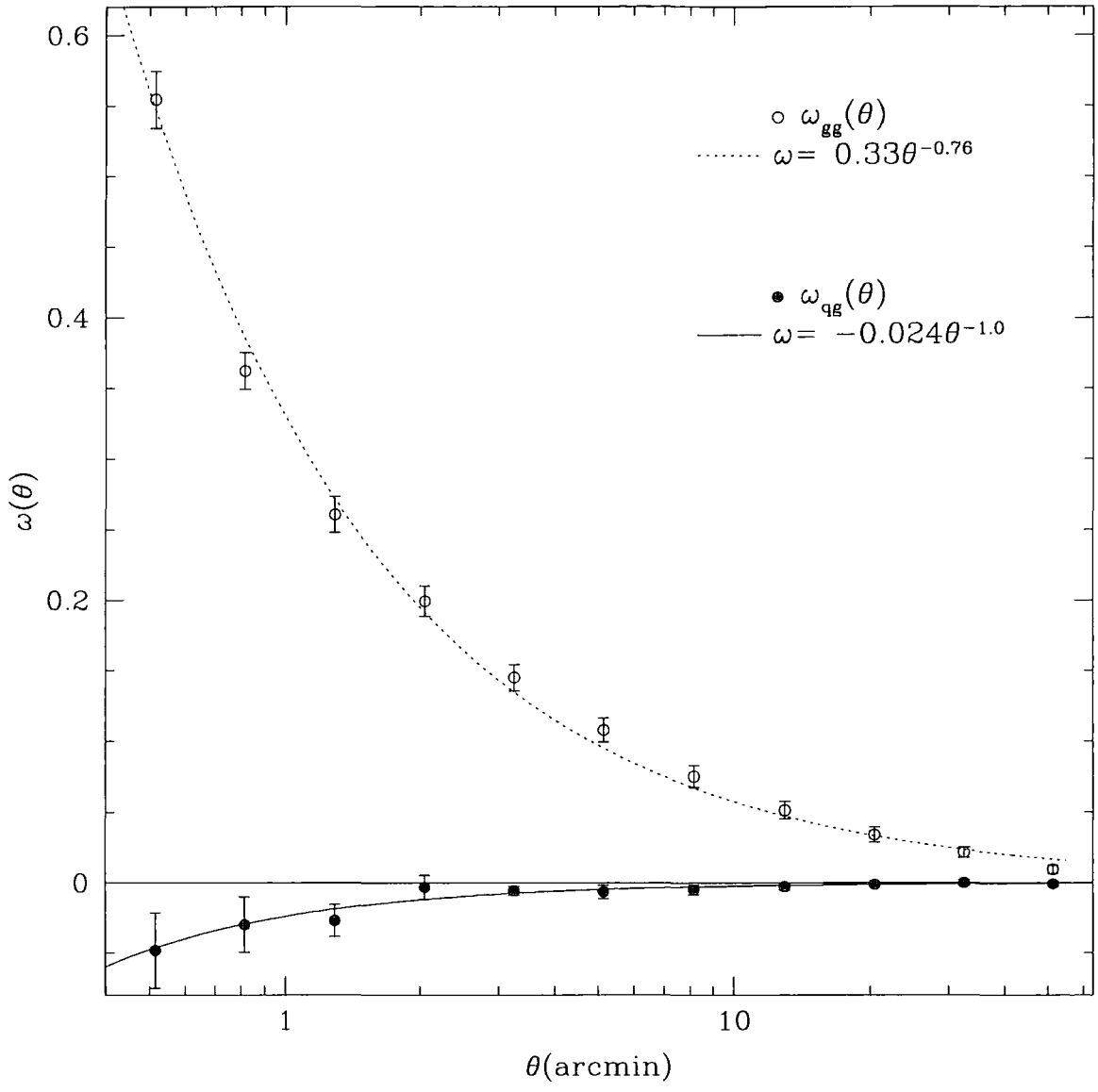


Figure 5.14: The galaxy-galaxy auto-correlation combined from SDSS EDR galaxies (in the region of the NGC) and APM galaxies (in the region of the SGC) is compared to the faint-end QSO-galaxy cross-correlation function. Both correlation functions are fitted by power laws, displayed by lines drawn through the respective data points. Error bars represent $1\sigma$ jackknife errors.

| | $\Omega_m = 0.3, \Omega_\Lambda = 0.7$ | | $\Omega_m = 0.3$ | | $\Omega_m = 1$ | |
|---|---|---|---|---|---|---|
| $z_{max}$ | 0.3 | 0.4 | 0.3 | 0.4 | 0.3 | 0.4 |
| $\bar{\kappa}$ | 0.01703 | 0.02812 | 0.01369 | 0.02154 | 0.04166 | 0.06389 |

Table 5.1: Here we list the values of $\bar{\kappa}$ as measured from Equation 5.11 for some different cosmological assumptions. We assume $z_s = 1.5$ for the QSO distribution median redshift and a smooth (homogenous) universe.

| | $\Omega_m = 0.3, \Omega_\Lambda = 0.7$ | | $\Omega_m = 0.3$ | | $\Omega_m = 1$ | |
|---|---|---|---|---|---|---|
| $z_{max}$ | 0.3 | 0.4 | 0.3 | 0.4 | 0.3 | 0.4 |
| $b$ | $0.13\pm^{0.08}_{0.07}$ | $0.21\pm^{0.14}_{0.12}$ | $0.10\pm^{0.07}_{0.06}$ | $0.16\pm^{0.10}_{0.09}$ | $0.32\pm^{0.20}_{0.18}$ | $0.48\pm^{0.31}_{0.27}$ |

Table 5.2: Here we list the values of $b$ as measured from Equation 5.52 for different cosmological models. The term $\Sigma_{cr}(z, z_s)$ has been calculated using Equation 5.9, with $z_s = 1.5$ for the QSO distribution median redshift.

### 5.6.1 Linear Bias Model

We now use Equation 5.19 (Williams & Irwin 1998) to determine the galaxy bias parameter, $b$. In Table 5.1 we show the term $\bar{\kappa}$ for the lensing slab, as expressed in Equation 5.11, determined for Einstein-de-Sitter (henceforth EdS), ΛCDM and open cosmologies. As the exact redshift distribution of APM galaxies and SDSS EDR galaxies to $b_J < 20.5$ is not known, we estimate the extent of the redshift distribution using the magnitude-redshift selection function developed by Baugh & Efstathiou (1993) which provides an excellent fit (Maddox et al. 1996) to the Stromlo-APM Survey redshift distribution (Loveday et al. 1992b, Loveday et al. 1992a) and is displayed in Fig. 2.6. Just over 95 per cent of the galaxy distribution is included out to $z_{max} = 0.3$ rising to just over 99.5 per cent out to $z_{max} = 0.4$. We calculate $\bar{\kappa}$ using Equation 5.9.

Using Equation 5.19, together with our measured values for the correlation function and taking a faint end slope of $\alpha = 0.29 \pm 0.05$ as determined in Chapter 3, we determine the galaxy bias as

$$b = 7.56 \pm^{4.83}_{4.24} \bar{\kappa}\theta^{0.24\pm0.30} \tag{5.52}$$

In Table 5.2 we display the values of $b$ for the different cosmologies outlined in Table 5.1. Note that a scale-independent model of bias is favoured only marginally by the errors. There is a scale-dependent ΛCDM model that is *not rejected* by our analysis

$$b \sim 0.1 \left( \frac{0.1\,h^{-1}\mathrm{Mpc}}{r} \right)^{-0.5} \tag{5.53}$$

(where $r$ is equivalent to $\theta$ expressed in $h^{-1}$Mpc at the mean redshift of our galaxy data) that would fit most of the previous measurements of the galaxy bias parameter out to $200\,h^{-1}$Mpc (Lahav et al. 2002, Verde et al. 2002, Hoekstra et al. 2002, Gaztañaga 2003). These measurements were plotted in Fig. 5.1 and compellingly suggest that $b \sim 1$ for $r > 10\,h^{-1}$Mpc. However, as a caveat we note that the fitted forms of our correlation functions probably do not probe scales greater than about $3\,h^{-1}$Mpc (30 arcminutes), anyway.

### 5.6.2 General Bias Model

Substituting the power-law slopes from Equation 5.51 into Equation 5.47, we find slopes for the mass and bias correlation functions of

$$\gamma = 2.24 \pm 0.43 \tag{5.54}$$

$$\gamma_b = -0.24 \pm 0.30 \tag{5.55}$$

We use these values in Equations 5.43 and 5.45 to determine the expected strength of $A_{gg}$ and $A_{gq}$ for a variety of cosmological assumptions, which we display in Table 5.3. The weight functions have been calculated as per Equation 5.27 using the redshift distributions for our QSO and galaxy samples. The QSO redshift distribution is measured from the 2QZ sample. The galaxy redshift distribution uses the redshift-magnitude relation from (Baugh & Efstathiou 1993) sampled to $b_J < 20.5$, the limiting magnitude of our galaxy samples. Both of these redshift distributions have already been plotted in Chapter 3. When calculating the lensing efficiency (Equation 5.41) we use a slope of $\alpha$=0.29, as calculated for the 2QZ faint-end number-magnitude counts in Chapter 3. Two things should be highlighted that arise from Table 5.3. Firstly, as the lensing efficiency is directly dependent on $\Omega_m$, the expected amplitude of QSO-galaxy cross-correlations is much higher in an EdS cosmology than for ΛCDM. Secondly, it seems that the QSO-galaxy cross-correlation amplitude is directly proportional to the square of Hubble's Constant. This is *not* a general feature of the model but arises simply because $\gamma + \gamma_b = 2$ in our analysis.

We now use Equation 5.49 to determine the galaxy bias and mass fluctuations in a ΛCDM cosmology, with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, for $0.1\,h^{-1}$Mpc scales. The result can

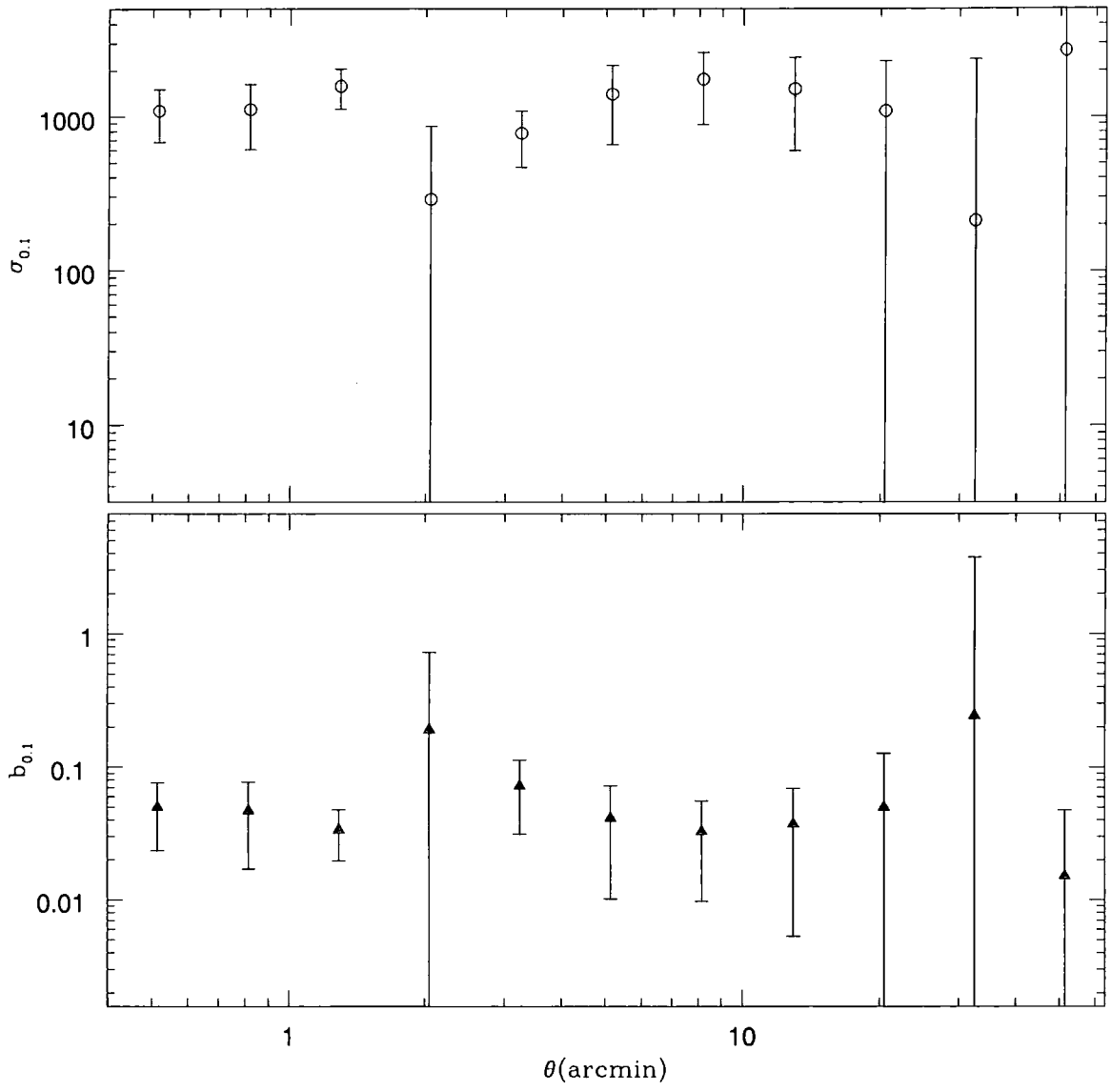


Figure 5.15: This figure presents measurements of galaxy bias and the amplitude of mass clustering on scales of $0.1\,h^{-1}$Mpc, determined across the range of the galaxy-galaxy auto-correlation and galaxy-QSO cross-correlation. Error bars are calculated from the ratio of the $1\sigma$ jackknife errors in each bin of the correlation functions.

| | $\Omega_m = 0.3, \Omega_\Lambda = 0.7$ | | | $\Omega_m = 1$ | | |
|---|---|---|---|---|---|---|
| $h$ | 0.5 | 0.7 | 1.0 | 0.5 | 0.7 | 1.0 |
| $A_{gg}(0.1\,\mathrm{Mpc}) \times 10^6$ | 0.128 | 0.231 | 0.432 | 0.150 | 0.270 | 0.507 |
| $A_{gq}(0.1\,\mathrm{Mpc}) \times 10^9$ | -0.0616 | -0.121 | -0.247 | -0.200 | -0.385 | -0.786 |
| $A_{gg}(0.2\,\mathrm{Mpc}) \times 10^6$ | 0.432 | 0.782 | 1.46 | 0.510 | 0.916 | 1.72 |
| $A_{gq}(0.2\,\mathrm{Mpc}) \times 10^9$ | -0.247 | -0.483 | -0.986 | -0.786 | -1.54 | -3.14 |

Table 5.3: Here we list the values of $A_{gg}$ and $A_{gq}$ as measured from Equations 5.43 and 5.45 for different cosmological models. Hubble's Constant is paramaterised as $H_0 = 100h\,\mathrm{kms}^{-1}\,\mathrm{Mpc}^{-1}$. $A_{gg}$ and $A_{gq}$ have been calculated assuming stable clustering evolution ($\epsilon = 0$ in Equation 5.26). We consider two scales for the measurements $0.1\,h^{-1}$Mpc, corresponding to an angular extent of 1 arcminute at the median redshift of our galaxy distribution and $0.2\,h^{-1}$Mpc, corresponding to the angular extent chosen by Gaztañaga (2003) in a similar analysis using a galaxy distribution with the same median redshift as our galaxy distribution.

be scaled to other models using Table 5.3 and Equation 5.49. In Fig. 5.15 we show the determination of the bias parameter, $b_{0.1}$ and mass fluctuation $\sigma_{0.1}$ on $0.1\,h^{-1}$Mpc scales, as determined across the range of our correlation functions. The variance-weighted mean values from this plot are

$$b_{0.1} = 0.04 \pm 0.01 \tag{5.56}$$

$$\sigma_{0.1} = 1013 \pm 206 \tag{5.57}$$

Note that the variance-weighted error-bar quoted is almost certainly too small - the method of variance-weighting assumes that the data are independent, which our data almost certainly aren't. We have used this method to obtain an error-bar for comparison, as this is the method advocated by Gaztañaga (2003). In our opinion, however, a fairer representation of the error would be estimated from the error on the data at about $1\,h^{-1}$Mpc (1 arcminute) in Fig. 5.15, which would give an error on $b_{0.1}$ of $\pm 0.12$ (and on $\sigma_{0.1}$ of $\pm 500$), in better agreement with the estimate noted in Table 5.2 for a $\Lambda$CDM cosmology, derived from the model of (Williams & Irwin 1998).

Had we instead performed the analysis for scales of $0.2\,h^{-1}$Mpc, we would recover

$$b_{0.2} = 0.05 \pm 0.01(0.04) \tag{5.58}$$

$$\sigma_{0.2} = 467 \pm 95(230) \tag{5.59}$$

Where the listed errors are the variance-weighted estimates and the values in brackets are simply taken from the error bar on the data at about $1\,h^{-1}$Mpc (1 arcminute) in Fig. 5.15. This is in good agreement with the result of Gaztañaga (2003), especially when we interpret the error-bars on the measurements of Gaztañaga (2003) as under-estimates, given that they assume the data in each bin are independent.

### 5.6.3 Discussion

In the main, the intent of this chapter was to study the cross-correlation between galaxies and a QSO sample drawn from the faint end of the QSO number-magnitude counts, to investigate whether the implications of such correlations for galaxy bias were consistent with the work of a few other authors working with bright QSO samples. A number of other cosmological implications arise from the model of (Gaztañaga 2003), though, and not discussing them would leave conclusions from this chapter incomplete. However, as our results largely confirm the work of Gaztañaga (2003), most of the cosmological implications of the model parameters deduced in this chapter are already discussed in depth by that author, so before discussing our measurement of bias we will only briefly outline the main cosmological predictions that arises from our statistical lensing analysis.

We determine the slope of the matter correlation to be $\gamma = 2.24 \pm 0.43$. Typical values for the slope of the matter correlation function found from N-body simulations (Hamilton et al. 1991, Peacock & Dodds 1996, Cooray & Sheth 2002) are 1 in the non-linear regime (0.5 at larger scales, 1.5 at the transition between linear and non-linear scales). Recently, Smith et al. (2003) have investigated small-scale non-linear correlations using a model particularly appropriate in this regime. Fig. 6 of Smith et al. (2003) allows for correlation slopes of $\gamma > 1.8$ for power spectrum slopes of $n > 0$. Recent data have probed power spectrum slopes on Megaparsec scales (Percival et al. 2001, Dodelson et al. 2002, Outram et al. 2003) and all are found to be negative. This means that a major prediction of our analysis is that the slope of the power spectrum will turn over and become positive at smaller scales.

Traditionally, galaxy bias is measured from the ratio between the galaxy auto-correlation and the mass auto-correlation $b^2(R) = \frac{\bar{\xi}_{gg(R)}}{\bar{\xi}(R)}$ (see Appendix B). The amplitude (i.e. $\sigma_{0.1}$) of the real-space galaxy auto-correlation function measured from the APM galaxy catalogue is well-constrained, even on scales as small as $0.1\,h^{-1}$Mpc (Baugh, Cole & Frenk 1996). The form of the mass correlation in a ΛCDM universe can be measured on the same scales very accurately from simulations (see Colin et al. 1999, or Fig. 11 of

Gaztañaga 1995, or Fig. 3 of Berlind & Weinberg 2002). Taken in ratio, the amplitude of mass and galaxy auto-correlations yield a value of $b \sim 0.6$. Thus, a value of $b \sim 0.1$ on these scales must be considered at odds with the standard prediction of ΛCDM. It turns out (see Gaztañaga 2003) that a positive slope for the small-scale mass power spectrum (as discussed in the previous paragraph) returns much larger amplitudes for the mass auto-correlation, consistent with $b \sim 0.1$. We refer the reader to Gaztañaga (2003) for the detailed argument. If, in future, this turnover in the power spectrum turns out not to be a viable proposition, we are left with a handful of possible alternatives. It could be the case that the ΛCDM hypothesis is simply wrong, which does not seem likely given the accumulating observational evidence in its favour (see Chapter 1). Our assumption of stable clustering ($\epsilon = 0$ in Equation 5.26) may not be valid at $z < 0.5$. The QSO number-counts might flatten in extremis at faint magnitudes. Or, there may be some untested systematic in our measurement of $\omega_{gq}$.

It is remarkable and compelling that our bias predictions agree so well with the predictions of other authors (Williams & Irwin 1998, Gaztañaga 2003) working with bright QSOs and (mostly) independent samples of QSOs and galaxies. It is difficult to imagine a systematic (other than statistical lensing) that could mimic both an anti-correlation between galaxies and faint QSOs and a positive correlation between galaxies and bright QSOs. Both of the theoretical systems we have used in this chapter, which utilise quite different lensing models to determine the amplitude of the galaxy bias parameter, $b$, consistently agree that on scales of 0.1 $h^{-1}$Mpc, galaxies are strongly anti-biased (i.e. $b << 1$) with a bias parameter of $b \sim 0.1$. Using a simple model (Williams & Irwin 1998) and assuming that lensing matter can be represented by a uniform slab extending out to the redshift where 95 per cent of the surveyed galaxies are included, and the lensed QSOs may be represented by a single population at $z = 1.5$, we find $b = 0.13 \pm^{0.08}_{0.07}$ for a ΛCDM cosmology. A more realistic model that traces spherical fluctuations in the underlying lensing matter on $0.1\,h^{-1}$Mpc scales, across the redshift distributions of our QSO and galaxy samples, yields $b = 0.04 \pm 0.12$. Although we cannot rule out a linear bias parameter on scales of $0.1\,h^{-1}$Mpc from the slopes of our fitted correlation functions, the fact that many other authors (see Fig. 5.1) significantly predict $b \sim 1$ on Megaparsec scales where we rule out $b \sim 1$ with strong significance in a variety of cosmologies (and certainly in ΛCDM ) suggests that bias is a strong function of scale on $100\,h^{-1}$kpc scales. It might be argued from Table 5.2 that in an EdS cosmology, $b = 1$ is only marginally ruled out when the lensing slab is extended to $z = 0.4$. However, the choice of the extent of the

slab is fairly arbitrary. In fact, if the bias is predicted by scaling from the amplitudes in Table 5.3, we would determine $b = 0.2 \pm 0.2$, with $b \sim 1$ on these scales strongly ruled out. We determine the slope of the scale-dependent bias parameter as $\gamma_b = -0.24 \pm 0.3$, which allows a model that could yield $b \sim 1$ on Megaparsec scales, so our results for galaxy bias are not particularly at odds with the recent work of other authors (see Fig. 5.1).

In the Introduction, we discussed the formalism that has emerged over the last decade to model galaxy-qso cross-correlations via N-body simulations (Bartelmann 1995, Dolag & Bartelmann 1997, Sanz, Martínez-González & Benítez 1997). In these simulations, an appropriate bias prescription (and often a matter correlation function slope) must be chosen to relate the simulation to the observed galaxy distribution. If, as seems to be the case, the bias is highly scale-dependent over the scales that weak lensing statistics typically probe, great care must be taken in constructing cosmological statistical lensing simulations, if predictions are to be made about the behaviour of lensing galaxies on these scales, rather than just lensing mass.

## 5.7 Conclusions

In this chapter, we have studied the cross-correlation of galaxies and background QSOs. Galaxies were drawn from the APM Survey in the region of the 2dF QSO Redshift Survey Southern Galactic Cap strip and the Sloan Digital Sky Survey Early Data Release in the Northern Galactic Cap strip. The QSO-galaxy cross-correlation function $\omega_{qg}$ suggests that QSOs and galaxies are anti-correlated with significance ($2.8\sigma$) and strength $-0.007$, over the angular range out to 10 arcminutes ($1\,h^{-1}$Mpc at the median redshift of our galaxy samples). This result is unique in that it is possibly the first significant anti-correlation detected between QSOs and galaxies that is not subject to small sample sizes or problems in categorically identifying the QSO population. We have also detected an anti-correlation between White Dwarfs and background galaxies and suggested a possible explanation of dusty Oort clouds (Oort 1950) around White Dwarfs.

We have outlined a number of possible explanations for the anti-correlation between QSOs and galaxies. Firstly, errors and the correlation estimator were proven robust against Monte Carlo simulations. Creating random catalogues that correct for dust in our Galaxy have no appreciable effect on the anti-correlation signal. Care was taken to demonstrate that there is no significant correlation ($0.5\sigma$ positive correlation) between stars (which were selected within the 2dF Survey in the same way as the QSO sample)

and our galaxy samples. We are thus confident that the anti-correlation between QSOs and galaxies is not a systematic error.

The colours of QSOs around galaxies were used to place strong limits on the effect dust (modelled with an absorption law appropriate to the Milky Way) could have in producing an anti-correlation of QSOs with galaxies. We do not rule out the possibility that some atypical dust distribution, that is inconsistent with dust observed in our Galaxy could explain the anti-correlation, nor do we deny that dust could account for as much as a third of the anti-correlation signal (at $2\sigma$ significance).

Aspects of our galaxy-QSO anti-correlation are consistent with the predictions of statistical lensing theory. The anti-correlation becomes more positive as we move towards brighter samples of QSOs, although not significantly. The signature of double magnification bias is hinted at in the radio-loud QSO sample. With no other explanation for the anti-correlation, we consider the cosmological implications if it is caused by statistical lensing. We consider two models that use quite different descriptions of the lensing matter and find they yield consistent predictions for the strength of galaxy bias on $0.1\,h^{-1}$Mpc scales of $b \sim 0.1 \pm 0.1$ (for a $\Lambda$CDM cosmology). The inferred strength of this result is in agreement with the work of Williams & Irwin (1998) and Gaztañaga (2003), although Williams & Irwin (1998) probed larger scales. The fact that Williams & Irwin (1998) found the same result but at larger scales is not necessarily a discrepancy, as they considered a particularly red galaxy population, dominated by early-type galaxies, and bias may well be a strong function of galaxy colour and type (Willmer, da Costa & Pellegrini 1998), although probably not sufficiently strong to explain the discrepancy (Norberg et al. 2002).

What is unique about our result, that galaxies are highly anti-biased on small scales, when combined with the work of Gaztañaga (2003), is that a consistent picture emerges that spans faint and bright QSO samples, which probe different parts of the QSO number-magnitude counts, where the counts have very different slopes. It is particularly compelling that agreement is reached between our result, based on an anti-correlation, and a result based on a positive correlation (Gaztañaga 2003), as it is difficult to imagine a systematic effect (other than statistical lensing) that can mimic both positive correlations and anti-correlations. We also confirm the measurement of $\gamma > 1.8$ ($1\sigma$ level) for the slope of the lensing matter cross-correlation made by Gaztañaga (2003). The slope of our power-law fits to the QSO-galaxy cross-correlation does not allow us to rule out a linear bias parameter but when we compare our measurement of $b \sim 0.1$ on $100\,h^{-1}$kpc scales to independent methods that determine $b \sim 1$ on $h^{-1}$Mpc scales (Hoekstra et al. 2002, La-

hav et al. 2002, Verde et al. 2002), we must conclude that bias is scale-dependent. Unfortunately, we cannot directly probe $h^{-1}$Mpc scales with our current method, as the QSO-galaxy anti-correlation is no longer significant on these scales.

As it increases in size, the Sloan Digital Sky Survey (Stoughton et al. 2002) should be best equipped to measure galaxy bias from the cross-correlation between QSOs and galaxies. As the SDSS contains a distribution of QSOs that mainly lie where the slope of the QSO number-magnitude counts steeply deviate from 0.4 ($\alpha \sim 0.75$) as compared to the fainter 2dF QSO sample ($\alpha \sim 0.3$), the expected lensing effect on SDSS QSOs is stronger and more significant effects will result. This has already been demonstrated by Gaztañaga (2003), who found a $4\sigma$ positive correlation between SDSS QSOs and galaxies. Further, the completed SDSS is more likely to uncover significant QSO-galaxy correlations out to larger scales. The SDSS Early Data Release comprises about 10 per cent of the total survey, so points which are marginally significant at the $1\sigma$ level in the EDR should be highly significant ($\sim 4\sigma$) in the full survey. This should allow galaxy bias to be probed on scales similar to scales probed by weak lensing shear methods (Hoekstra et al. 2002) and may just approach scales probed by galaxy clustering methods (Lahav et al. 2002, Verde et al. 2002). However, the boldest evidence that a signal genuinely results from the statistical lensing of QSOs will remain a consistent effect across the range of the QSO number-magnitude counts, meaning it is important to confirm findings in bright QSO samples by considering the same processes in faint QSO samples, as we have done in this chapter. With this in mind, future large surveys of faint samples of QSOs will remain an important probe of how galaxies are biased relative to the underlying lensing matter.

# Chapter 6

# *The Faint-End Slope of the QSO Number-Magnitude Counts*

## 6.1 Introduction

In previous chapters, we have studied the cross-correlation of $b_J < 20.85$ 2dF QSO Redshift Survey QSOs with various foreground objects, modelling the resulting anti-correlations in terms of statistical lensing. To make predictions based on the magnification of QSOs by gravitational lensing requires a good understanding of how QSO numbers evolve with magnitude and, particularly, knowledge of the slope of the integrated QSO number-magnitude counts. Repeatedly in this thesis, we have used a value for the faint-end slope of the QSO number-counts based on extending a fit to 2dF QSO counts at $b_J < 20.85$ to fainter magnitudes, also taking into account data from two surveys to $b_J \lesssim 22$, which together cover about 1 square degree of sky (Koo & Kron 1988, Boyle, Jones & Shanks 1991). However, most of the known QSOs surveyed to $b_J > 21$ are from surveys carried out over a decade ago, a decade in which the technology necessary to carry out spectroscopic surveys should have progressed to an extent that QSOs can be identified to $b_J > 21$ at a fraction of the observational overhead (as multi-object fibre-optic spectrographs now typically have 2-5 times as many fibres, as compared to a decade ago).

The first wide-ranging QSO surveys were carried out at radio frequencies, e.g., Schmidt (1968), although it was quickly realised that radio-quiet QSOs had quite different properties to their radio-loud counterparts (Schmidt 1976) and merited separate surveys. The pioneering survey of radio-quiet QSOs was carried out by Braccesi, Formiggini & Gan-

dolfi (1970), who obtained photometry for 176 particularly blue objects to $b < 19.5$ in 36 square degrees of sky, arguing that as QSOs tended to have an ultraviolet excess (UVX) as compared to the main stellar locus, multicolour photometry could be used to define a QSO sample in the $(U-B)$,$(B-V)$ plane independently of the radio flux of the QSOs.

The work of Braccesi, Formiggini & Gandolfi (1970) remained the only major UVX QSO candidate sample for around a decade, at which point, the UVX technique started to become the standard method of selecting large numbers of candidate QSOs for spectroscopic follow-up, although the objective-prism technique provided an alternative methodology of finding higher redshift QSOs via their Lyman-$\alpha$ emission (Osmer & Smith 1980). As spectrographs improved, the number of confirmed UVX QSOs rose rapidly over the next decade. Schmidt (1983) selected ~1800 objects in ~10,000 square degrees, with $U-B < -0.44$ and after obtaining spectra for each object, found 114 QSOs to a limiting magnitude of $B = 16.16$. A year later, the first spectroscopically complete UVX sample was garnered, when 35 QSOs with $B < 19.80$ were specified over a region of 1.72 deg$^2$ (Marshall et al. 1984). Koo, Kron & Cudworth (1986) carried out the first deep UVX QSO survey to $B \sim 22.5$ (complete to $B \sim 21$), spectroscopically identifying 30 QSOs in 0.3 square degrees of sky (Koo & Kron 1988). A similar survey to $b_J \sim 22$ by Marano, Zamorani & Zitelli (1988) obtained spectra of 23 QSOs in 0.69 deg$^2$. The first single survey to measure more than 100 high-quality QSO spectra was probably that of Crampton, Cowley & Hartwick (1987), who obtained spectra for 163 QSOs to $B = 20.5$, a feat surpassed by Boyle, Fong Shanks & Peterson, who identified 420 QSOs to $B \lesssim 21$ in a complete survey of 1400 UVX objects.

Two practically complete samples of QSOs to a faint limit of $B \lesssim 22$ exist in the literature (Boyle, Jones & Shanks 1991, Zitelli et al. 1992). Combined, these surveys contain about 80 QSOs fainter than $b_J = 20.5$. Perhaps surprisingly, when we consider that the typical multi-fibre spectrograph can observe a multitude of objects simultaneously, as compared to a decade ago, the survey of Boyle, Jones & Shanks (1991) is still the largest QSO survey to $B \lesssim 22$. Much QSO survey work in the last ten years has either centred on finding QSOs at large redshift (Warren, Hewett & Osmer 1994, Storrie-Lombardi et al. 2001, Sharp et al. 2001, Fan et al. 2001, Monier et al. 2002) or finding small samples of very faint ($B \sim 23-25$) QSOs, with the notable exception of the work of Hall et al. (1996), extended by Kennefick et al. (1997), a survey of 0.83 deg$^2$ that contains 55 QSOs in total, 25 or so of which are fainter than $B \sim 21$. Of course, the 2dF QSO Redshift Survey (Croom et al. 2003) contains a huge number of QSOs compared to its predecessors

($\sim 23,000$ QSOs), and the Sloan Digital Sky Survey QSO catalogue, when completed, is expected to be yet another factor of 4 larger (Schneider et al. 2002). However, neither of these surveys were designed to observe objects as faint as $b_J = 21$ and if the Sloan Survey discovers a large number of QSOs that *are* so faint, it will only occur as a serendipitous catalogue representing a few per cent of a large number of objects.

There is certainly justification for a survey of QSOs to magnitudes fainter than $b_J \sim 21$, from our perspective to determine the faint-end slope of the QSO number-magnitude counts, although other similar analyses also benefit from faint QSO data, notably studies of the evolution of the QSO luminosity function (the quantity of QSOs of given *absolute* magnitude) with redshift, e.g., Boyle et al. (2000) and references therein. In Early 2002, we were granted 4 nights of dark time on the 4.2m William Herschel Telescope, one of the Isaac Newton Group of telescopes and a further 3 nights of service time on the 3.9m Anglo-Australian Telescope (alongside two other projects) to study the form of the QSO number-counts in the magnitude range $20.5 \leq b_J < 22.5$.

This chapter concerns the reduction and preliminary analysis of two faint QSO surveys. In Section 6.2 of this chapter we discuss our selection of candidate QSOs from the Sloan Digital Sky Survey Early Data Release. in Section 6.3 we outline our spectroscopic observation of candidate QSOs and our data reduction. In Section 6.4 we discuss the completeness of our QSO surveys, derive the number-magnitude counts of our spectroscopically confirmed QSOs and discuss the extension of our concept of using Sloan Digital Sky Survey photometry to extend the 2dF QSO Survey to fainter magnitudes. In Section 6.5 we summarise what is discussed in this chapter.

## 6.2 Selection of QSO candidates

Our target catalogue of QSO candidates was selected from the Sloan Digital Sky Survey (henceforth SDSS) Early Data Release (Stoughton et al. 2002). As a first cut in our data sample, only objects classified as a star by the SDSS data reduction pipeline were considered. Broadly, the SDSS distinguishes stars from galaxies by denoting an object as non-stellar if it has a magnitude obtained by fitting a galaxy intensity profile that deviates more than 15 per cent from the magnitude obtained by fitting a point-spread function (henceforth PSF). In total, the Early Data Release (henceforth EDR) of the SDSS supplies photometry for nearly 2.75million stellar objects.

### 6.2.1 SDSS and UKST Magnitude Systems

To compare the magnitudes of objects detected in both the 2dF QSO Redshift Survey (henceforth 2QZ) and the SDSS EDR, and ultimately select objects from the SDSS that obey the 2QZ QSO targeting criteria, required a conversion between the SDSS EDR photometric system (consisting of 5 wide band filters designated $u', g', r', i', z'$) and the United Kingdom Scmidt Telescope (henceforth UKST) bands ($u, b_J, r$) used in the APM survey, the parent catalogue of the 2QZ. Blair & Gilmore (1982) quote the conversion between the UKST system and the standard Johnson-Cousins photometric passbands as

$$u = U - (0.03 \pm 0.05) \times (U - B) \quad (6.1)$$

$$b_J = B - (0.28 \pm 0.04) \times (B - V) \quad (6.2)$$

$$r = R - (0.00 \pm 0.05) \times (R - I) \quad (6.3)$$

where $U, B, V, R, I$ denote Johnson-Cousins bands. Note that the 2QZ adopted $u = U$ in preparing their photometry. However, standard star work carried out in the preparation of the 2QZ input catalogue determined a zero-point offset leading to the u-magnitude adopted by the 2QZ actually being $u^* \simeq u - 0.24$ (R. J. Smith, *private communication*), which is the $u$ band relationship we will adopt.

Fukugita et al. (1996) supply transformations from the SDSS bands to Johnson-Cousins bands based on modelling the expected SDSS colours for 175 standard stars

$$g' = V + 0.56(B - V) - 0.12 \quad (6.4)$$

$$r' = V - 0.49(B - V) + 0.11 \quad (6.5)$$

$$r' = V - 0.84(V - R) + 0.13 \quad (6.6)$$

$$u' - g' = 1.38(U - B) + 1.14 \quad (6.7)$$

$$g' - r' = 1.05(B - V) - 0.23 \quad (6.8)$$

for which they estimate errors in the range 0.01 to 0.07. Ongoing measurements of standard stars (Smith et al. 2002) suggest typical systematic errors of about $\pm 0.05$ arise in the transformations of Fukugita et al. (1996) for colours between 1 and -1 that are measured from magnitudes in adjacent passbands - i.e., $U - B$, $B - V$, $V - R$. These errors of $\pm 0.05$ will dominate the measurement by Smith et al. (2002) for the internal consistency of the SDSS magnitudes of (at most) 0.02.

Combining all of the above Equations yield the following transformations

$$
\begin{aligned}
(u^* &= u - 0.24) && (6.9)\\
u &= u' - 0.28(u' - g') + 0.42(g' - r') - 0.61 && (6.10)\\
b_J &= g' + 0.16 + 0.15(g' - r') && (6.11)\\
r &= r' - 0.16 - 0.09(g' - r') && (6.12)\\
& && (6.13)
\end{aligned}
$$

The errors associated with the colour coefficients in these transformations lead to an estimate of $\pm 0.1$ in $u$ and $r$ and $\pm 0.08$ in $b_J$. This means that the error associated with $u - b_J$ and $b_J - r$ colours arising from the transformation of the magnitudes of SDSS EDR objects from SDSS to UKST bands is $\pm 0.13$. We have adopted an error of $\pm 0.05$ for each of the Fukugita et al. (1996) transformations.

The stellar objects that we had taken from the SDSS EDR had their (PSF) magnitudes converted to the UKST system as outlined above. As a check of our transformation between the SDSS and UKST magnitude systems, we determined the relative magnitudes of objects in the SDSS EDR and 2QZ that lie within 2 arcseconds of each other. In Fig. 6.1, we plot the difference between derived SDSS $b_J$ magnitude and 2QZ $b_J$ magnitude for every star identified in the 2QZ (with "11" spectral quality) that lies within 2 arcseconds of an SDSS object. We plot these differences against the derived SDSS $b_J$ magnitude. The mean and standard deviation is calculated for $18.25 \leq b_J(\mathrm{SDSS}) < 20.85$ and displayed in Fig. 6.1 as the solid line and errorbar. In calculating these statistics, we ignore objects that have a difference in $b_J$ between the SDSS and 2QZ of more than 0.5 magnitudes, a cut which discards 1 per cent of all objects.

If we consider *every* 2QZ "11" star that has a match within 2 arcseconds in the SDSS EDR, the mean difference between the SDSS and 2QZ $b_J$ magnitudes is 0.00, with a standard deviation of 0.14. Note that the mean and median are in good agreement. The difference between our estimated error (of $\pm 0.08$) in $b_J$ that would arise purely from transforming magnitudes from SDSS to UKST bands and the error illustrated in Fig.. 6.1 of ($\pm 0.14$), will arise due to measurement errors in the 2QZ and SDSS EDR magnitudes. If each magnitude in the SDSS EDR and 2QZ are, on average, accurate to about 0.07 magnitudes, then the resulting total standard deviation (based on assuming the colour transformation from SDSS to UKST magnitudes introduces 0.08 magnitudes of error into the SDSS value of $b_J$) would be 0.14. Note that this would mean that the dominant

error in our SDSS $b_J$ measurement arises from the magnitudes themselves, not from the transformation between systems.

The methodology used to calculate the zeropoint offset (mean) and error (standard deviation) in comparing SDSS and 2QZ $b_J$ magnitudes can similarly be used to estimate errors in SDSS $u - b_J$ and $b_J - r$ colours. We find a mean offset of $0.26 \pm 0.25$ for $u - b_J$ and $0.04 \pm 0.25$ for $b_J - r$, both in good agreement with their expected zeropoints, given that the $u$ zeropoint is actually $u^* \simeq u - 0.24$, as professed in Equation 6.9. We find no trend in zeropoint (and error) with $b_J$ magnitude in any of this analysis, except in the brightest and faintest bins.

### 6.2.2 UVX Selection of QSO Candidates

The 2QZ distinguishes QSOs from Main Sequence stars using the UVX selection method, a simple cut in colour-colour space. Braccesi, Formiggini & Gandolfi (1970) illustrate at some length why QSOs lie bluewards of Main Sequence stars and why the UVX selection method works. Véron (1983) estimate the UVX method returns a QSO sample that is 90 per cent complete in the redshift range $z < 2.2$ although some estimates suggest that more like a third of QSOs are missed by UVX selection (Graham, Clowes & Campusano 1999, Rengstorf et al. 2001). As only UVX stellar objects are targeted in the 2QZ, QSOs at lower redshifts, which lie in galaxies that appear extended (and are reddened by these host galaxies) will be missed by the 2QZ survey, consequently, the 2QZ exhibits large-scale homogeneity in the redshift range $0.3 < z < 2.2$ (Outram et al. 2003). The UVX selection criteria used by the 2QZ are

$$u^* - b_J \leq -0.36 \tag{6.14}$$

$$u^* - b_J < 0.12 - 0.8(b_J - r) \tag{6.15}$$

$$b_J - r < 0.05 \tag{6.16}$$

These colour cuts were chosen to balance the contamination of QSO candidates by stars against the completeness of the QSO sample itself. As we wish our faint sample of candidates to have similar completeness properties to the 2QZ, we adopt the 2QZ colour selection criterion.

In Fig. 6.2, we compare $u - b_J$ against $b_J - r$ colours for identified QSOs and stars from the 2QZ (with "11" spectral quality) and their counterparts from the SDSS EDR. 2QZ and SDSS EDR objects are considered to be the same object if they lie within

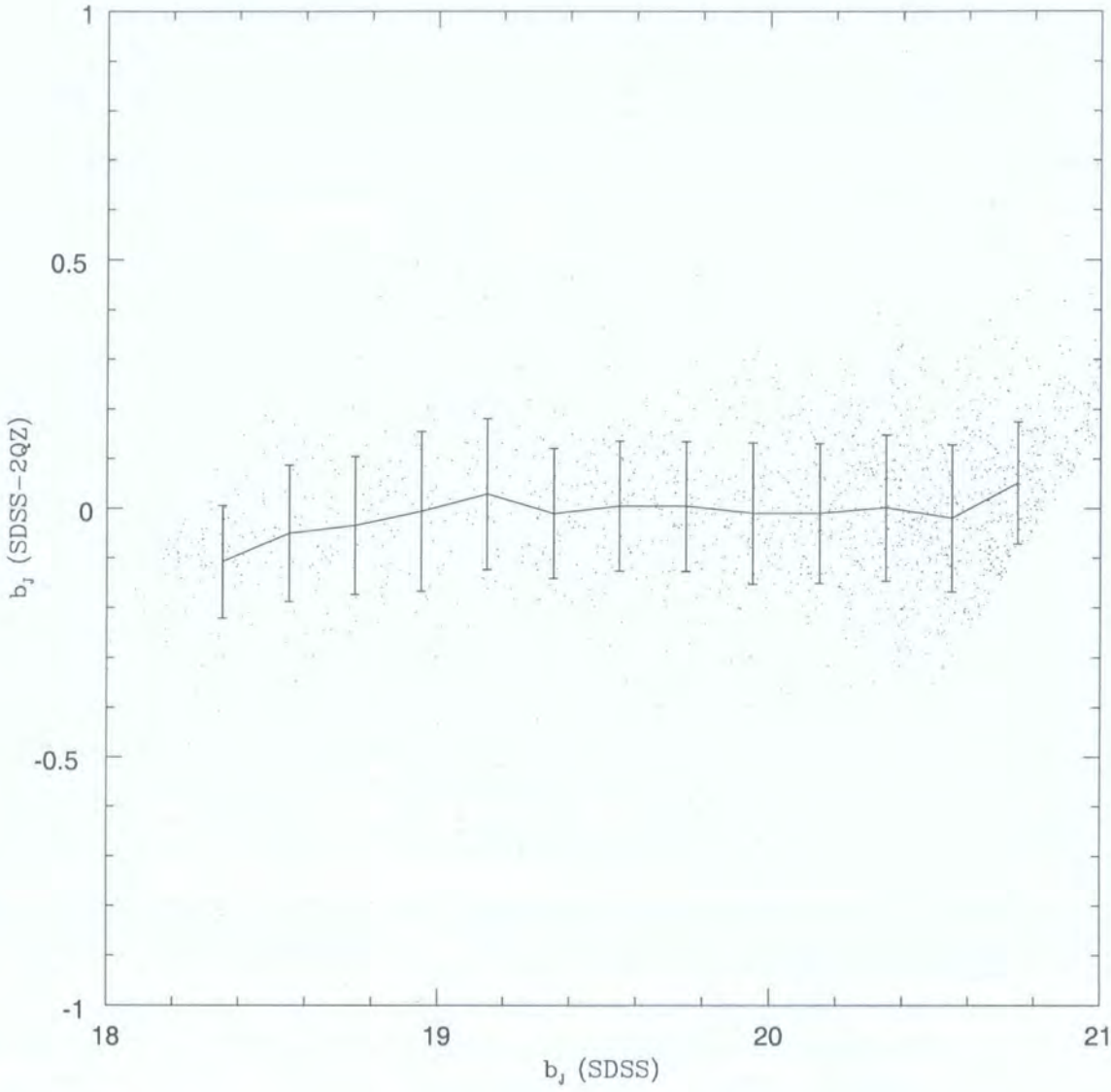


Figure 6.1: The offset between the UKST $b_J$ band and the SDSS equivalent, where the conversion between SDSS $g'$ and $r'$ bands and UKST $b_J$ band is made using Equation 6.12. Each dot represents a comparison between a "11" spectral quality 2QZ star and an SDSS EDR object that lies within 2 arcseconds of it. The line marks the mean value of the difference between SDSS and 2QZ $b_J$ magnitude across the range of SDSS $b_J$ magnitude. Error-bars are the standard deviation in that bin. In calculating the mean and standard deviation, a cut has been made excluding SDSS to 2QZ magnitude comparisons that deviate more than 0.5 in the $b_J$ band. This rejects just over 1 per cent of objects (i.e. it is a $3\sigma$ cut).

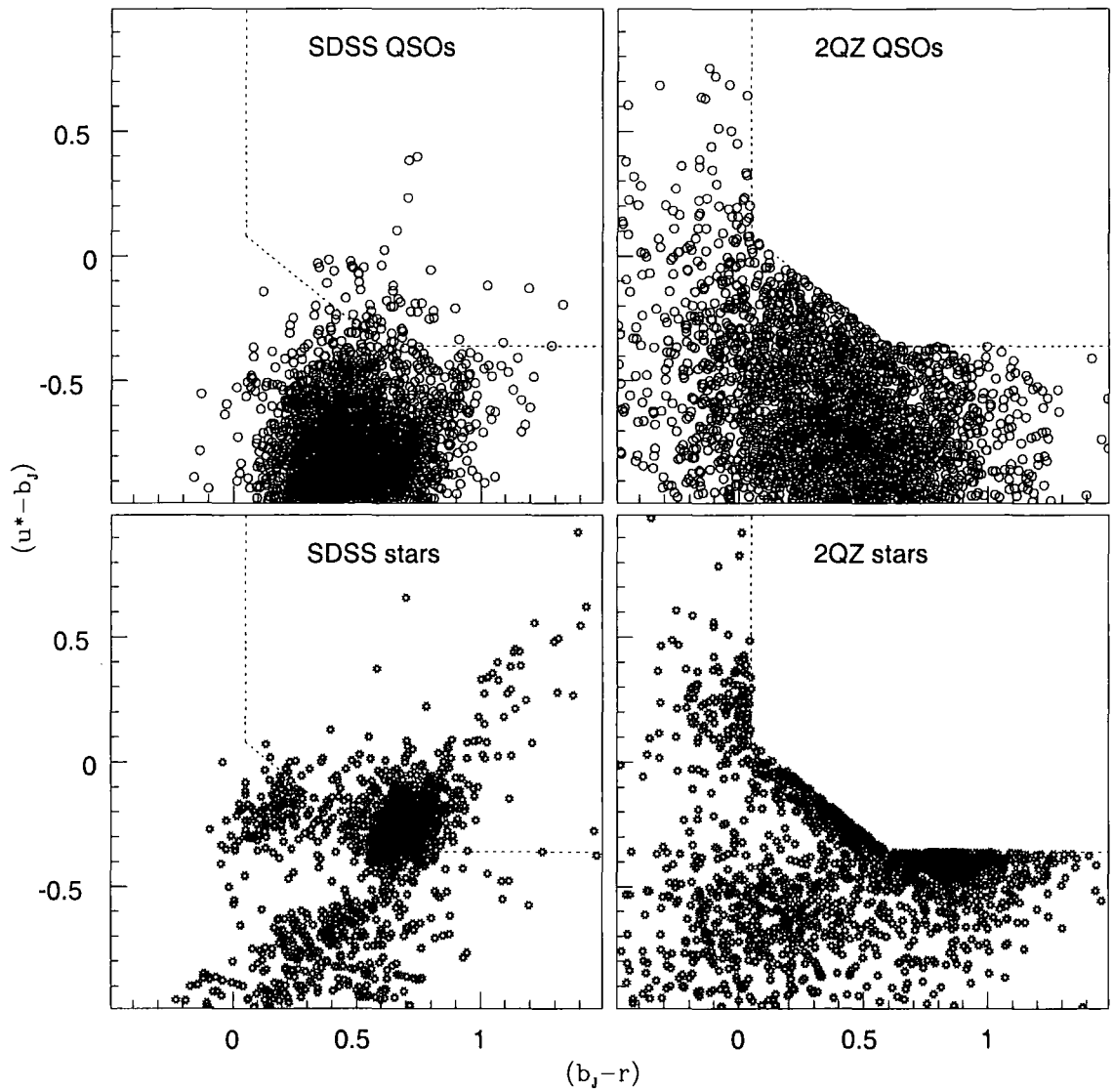


Figure 6.2: 2QZ colour cuts by object type. Panels on the right-hand side show the $(u - b_J, b_J - r)$ plane for ("11" spectral identification) 2QZ objects in the region of the SDSS EDR, with stars represented by stars and QSOs by circles. Panels on the left-hand side show "the same" objects (where SDSS objects are considered matched to their 2QZ counterparts if they lie within 2 arcseconds of them), with magnitudes as measured in the SDSS EDR but converted to UKST bands $u^*, b_J$ and $r$. The dotted lines represent the colour-cuts used to define the 2QZ input catalogue (Smith et al. 1997).

2 arcseconds of each other. SDSS magnitudes are transformed into the UKST system using Equations 6.9 through 6.13. The dashed lines show the colour criteria used to define the 2QZ input catalogue (Smith et al. 1997).

Over 98 per cent of QSOs identified in the 2QZ would be included in a catalogue based on the same colour cuts tranformed from SDSS EDR magnitudes. The QSO colours appear to be more tightly defined in the SDSS EDR - this might result from QSOs varying their magnitude between the 2QZ input catalogue $u, b_J$ and $r$ plate observations, which were often taken many years apart. More surprisingly, only 40 per cent of stars identified in the 2QZ would be included in a catalogue based on transformed SDSS EDR magnitudes. The reason for this appears to be that SDSS magnitudes are measured more accurately for these stars. This is not to say that the SDSS magnitudes are generally more accurate than the 2QZ - although the SDSS aim is a typical error of 0.03 in their colours (Smith et al. 2002), where the 2QZ colours might be in error by as much as 0.2 in the fainter bins (R. J. Smith, *private communication*) - rather, that the distribution of stars that made it into the 2QZ survey may be more likely to have poorly estimated colours than those that were rejected, whereas these same stars might have excellent photometry in the SDSS EDR. In any case, the $u - b_J$ against $b_J - r$ colour plane suggests that choosing QSO candidates using the typical 2QZ selection criteria should lead to a QSO fraction as good as the 2QZ (58 per cent of candidates turned out to be QSOs in the 2QZ), at least if completeness statistics for $20.5 \leq b_J < 22.5$, our target magnitude range, are similar to those for $18.25 \leq b_J < 20.85$. Note that all 35 QSOs fainter than $b_J = 20.85$ identified in the survey of Boyle, Jones & Shanks (1991) would be included in a target catalogue selected with the criteria outlined in this section.

## 6.3 Observations and Data Reduction

Having selected QSO candidates in the range $20.5 \leq b_J < 22.5$ from stellar objects in the SDSS EDR using the 2QZ UVX criteria, several fields containing these candidates were observed. From May 11th to May 14th 2002, 4 dark nights of time were allocated for us to observe our QSO candidates using the Wide Field Fibre Optical Spectrograph (WYFFOS) on the William Herschel Telescope (WHT). The observers listed for this project were Prof. T. Shanks and A. D. Myers. From March 14th to March 16th 2002, we were granted a portion of fibres (alongside a second project) over 3 dark nights to observe faint QSO candidates using the 2 Degree Field Spectrograph (2dF) on the Anglo

Australian Telescope (AAT). The observer listed for this project was Dr. P. J. Outram. In this section, we describe the spectrographic observation of QSO candidates with these two instruments and subsequent data reduction.

### 6.3.1 Observations with WYFFOS

The recently commisioned small-fibre module for AF2/WYFFOS (Corradi et al. 2001) provides 150 science fibres (as distinguished from the additional 10 fiducial fibre bundles, used to observe bright stars to guide the instrument) over a 1 degree (diameter) field. The fibres are 1.6 arcseconds in diameter on the plane of the sky, a diameter chosen to minimise the amount of sky observed around sources, while still being significantly larger than the typical median seeing at the WHT of $0.69 \pm 0.45$ arcseconds (Muñoz-Tuñòn, Varela & Mahoney 1998), minimising light-loss around the fibres. AF2 (Autofib2) is a robotic positioner that places fibres over the observing plate according to a user-defined configuration. Fibre-placement takes 30-50 minutes, depending on the complexity of the fibre configuration, and only one instrument is provided, meaning observations must cease while fibres are configured. Fibres are not permitted to cross, which generally means it is very difficult to assign more than 100 fibres to sources across the 1 degree field. The instrument suffers considerable vignetting in the field of view outside of a plate diameter of about 40 arcminutes, causing the source to be elongated in the field of view and, consequently, considerable light-loss by fibres placed more than 20 arcminutes from the plate centre. Ultimately, if 4-8 of the fiducial fibres are used and 5-15 science fibres are reserved to observe sky to subtract from the targeted sources, it is very difficult to observe more than 50 unvignetted sources at a sky density of 200-300 sources per square degree.

We selected QSO candidates in two sets of fields of 1 degree diameter. The first set of fields were in the SDSS EDR stripe centred at 0° Declination (with Right Ascension coverage corresponding to the 2QZ NGC strip), the second set of fields were assigned in the SDSS EDR strip at 59-60° Declination and 257-261° Right Ascension. Both sets of fields were chosen to have low absorption, according to the dust maps of Schlegel, Finkbeiner & Davis (1998) (i.e. absorption in the $b_J$ band of less than 0.1 magnitudes). The two different sets of fields allowed us to observe at low airmass throughout the night, as at the latitude of the WHT the first set of fields were at their closest approach to the zenith in the first half of the night, while the fields at higher declination were highest in the sky during the second half of the night. The largest airmass of any of our observations of target QSO fields was thus 1.3 (i.e., our fields were never observed below about 50°

altitude). In May at the WHT, the time between evening and morning twilight is about 7.5 hours. Over any night, our schedule was to observe a field at 0° Declination for 3 hours, wait 45 minutes while AF2 configured the fibre array to a second field, then observe a field at 60° declination. This left ample time to take bias frames, flat frames and arc frames throughout the night.

Our observing strategy consisted of taking a series of 1800 second exposures, with arcs (a series of 0.1, 2, 5, and 50 second exposures of a neon lamp) and flats taken in between each exposure for calibration purposes. Our basic flat field exposures consisted of 0.3 second exposures of a tungsten lamp, however, we also took 120 second exposures of "blank" fields by offsetting 10 arcseconds from our program field in a different direction after every two science exposures. We will refer to these "blank" fields as *offset sky exposures*. Our spectroscopy was carried out using an R300B grating, with a spectral range of 3540Å, allowing us to observe in the range 3700−7200Å, the minimum wavelength being sufficient to detect the broad MgII QSO emission line to a redshift of $z \sim 0.3$. Our CCD of choice was the $1024 \times 1024$ pixel TEK6 detector. SIGNAL, a software facility provided by Chris Benn at the Isaac Newton Group of Telescopes, suggested that a 3 hour exposure using WYFFOS, with this CCD and grating combination should be sufficient to detect objects to $B < 22.4$ in 1.3 arcsecond seeing and at an airmass of 1.3, to a signal-to-noise ratio of ∼8. A base exposure time of 1800 seconds is sufficient to ensure many more photons per Angstrom from objects with $B < 22.4$ than the expected number of electrons in readout noise from the CCD (i.e. our observations would not be limited by the noise from the CCD).

Fibres were arranged on each of our fields at random, so as to fairly sample QSOs in every bin of magnitude. Around 10 fibres were set aside for sky in each exposure, to facilitate the subtraction of the sky contribution from our derived spectra. A minimum of 4 fiducial bundles in each field were reserved to observe bright stars to guide WYFFOS during observation. Guide-stars were selected from the SDSS EDR in the magnitude range $14 < V < 15$ using colour transformations from $u'$ and $g'$ to the V band from Fukugita et al. (1996), and distributed around the field, to ensure fibres were accurately positioned on objects over the entire plate. In our 4 night run at the WHT, the equivalent of about 2.5 nights of good data were obtained, with the remaining time lost to a combination of cloudy weather and technical problems. A summary of our observations is listed in Table 6.1. We obtained our proposed total integration of 3 hours for 4 of the observed fields. Note that the $13^h44^m$ field was observed along with some X-ray QSO candidates

| Field Centre | | | | Seeing | Exposure | Fibre Assignments | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | $\delta$ | | (Or Weather) | Total | Candidates: Assigned/Total | | | |
| $h$ | $m$ | $^\circ$ | $'$ | $''$ | $s$ | $< 60'$ | $< 40'$ | Sky | Fiducial |
| 12 | 30 | 00 | 00 | $\sim 0.8$ | 10800 | 113/282 | 45/125 | 7 | 6 |
| 12 | 50 | 00 | 00 | $\sim 1$ | 10800 | 104/247 | 50/120 | 12 | 7 |
| 13 | 20 | 00 | 15 | (cloudy) | 3600 | 109/332 | 47/147 | 12 | 4 |
| 13 | 44 | -00 | 15 | $\sim 1.2$ | 10800 | 76/253* | 28/104* | 13 | 4 |
| 17 | 19 | 61 | 00 | $\sim 0.8$ | 10800 | 105/302 | 39/148 | 7 | 7 |
| 17 | 25 | 59 | 00 | (cloudy) | 7500 | 103/289 | 37/125 | 5 | 5 |
| 17 | 31 | 57 | 00 | $\sim 1.3$ | 9000 | 101/271 | 43/120 | 11 | 4 |

Table 6.1: Fields observed on the WHT with AF2/WYFFOS on the nights of the 10th, 11th, 12th and 13th of May 2002. The field centre is listed in J2000 coordinates. Four fields were observed to our desired exposure time of 3 hours. Along with the number of sky and fiducial fibres assigned in each field, we list the number of candidate QSOs observed over the number of candidate QSOs in the field. The column headed $< 60'$ lists this information for the entire ($1^\circ$ diameter) WYFFOS field. The column headed $< 40'$ is the same information within a radius of $20'$ of the field centre (i.e. the unvignetted field of view of WYFFOS). *The $13^h44^m$ field was observed simultaneously with a number of X-ray QSO candidates for a separate project, hence the number of available fibres for our QSO candidates was reduced.

for another project, hence the number of fibres available for UVX QSO candidates was reduced.

### 6.3.2 WYFFOS Data Reduction

Our spectra were reduced using the Image Reduction and Analysis Facility IRAF (e.g., Gilmore & Allen 1990) on the Durham Starlink Node. There is a dedicated package for reducing WYFFOS spectra, called *wyfred* but it suffers from a number of bugs and is currently being discontinued as support software for WYFFOS data reduction (Romano Corradi, *private communication*). Our data were thus reduced mainly using various sub-routines included in the *wyfred* package, particularly the *dofibers* software. Our CCD images were unpacked, a bias frame was used to subtract the CCD overscan region from the images, and a median bias frame was subtracted from each image. Cosmic rays were also clipped by inspecting each CCD image for bright pixels and interpolating across neighbouring pixels to remove pixels that were $10\sigma$ brighter than the median pixel value along the fibre. The positions of the fibres on the CCD frame were traced with polynomials - usually a linear fit was sufficient but up to a fourth order polynomial was necessary for fibres placed near the edges of the CCD. The mean residual of the fit to the the fibre positions was 2 per cent and the fit was always better than 5 per cent.

Once the fibre positions in each exposure were established (and each exposure was clipped and free from cosmic rays), the transmission of light through each fibre in our observational set-up was determined from our various flat fields. Each series of Tungsten flats that shared a fibre configuration were averaged into a single, mean lamp flat. The spectrum from each fibre of the mean lamp flat were averaged and smoothed to find the mean fibre throughput, then each individual spectrum was divided by the mean throughput to find the variation from the mean throughput on a fibre-by-fibre basis. We will call this the variation function. Each offset sky exposure that shared a fibre configuration was averaged and each spectrum in the exposure was smoothed and the mean number of counts fitted, to determine the transmission on a fibre-by-fibre basis for this particular fibre configuration. The transmission was then multiplied by the variation function and the resulting function for each fibre was divided by the mean across all fibres to create the fibre-by-fibre normalised response of the observing system.

We were now ready to reduce the actual QSO candidate spectra. Each (already traced) fibre position on a QSO candidate was divided by the fibre-by-fibre normalised response function for that particular fibre configuration. Next, we calibrated the wavelegth of the

| Field Centre | | | | | | | | | Exposure | Fibres |
|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | | $\delta$ | | | Area | Seeing | Airmass | Total | Candidates |
| $h$ | $m$ | $s$ | ° | ′ | ″ | *sq.deg.* | ″ | | $s$ | Assigned/Total |
| 10 | 34 | 57.34 | +01 | 23 | 16.6 | 1.55 | 1.2-1.3 | 1.3-2.0 | 21600 | 71/189 |
| 11 | 05 | 46.71 | -01 | 26 | 19.62 | 0.56 | 1.2-1.3 | 1.2-1.3 | 21600 | 37/99 |
| 12 | 54 | 18.12 | +00 | 03 | 08.7 | 3.14 | 1.3-1.5 | 1.2 | 21600 | 146/604 |
| 14 | 00 | 35.46 | -00 | 55 | 28.5 | 1.48 | 1.3-1.5 | 1.2-1.3 | 21900 | 84/273 |

Table 6.2: Fields observed on the AAT with the 2dF spectrograph on the nights of the 14th,15th and 16th of March 2002. The field centre is listed in B1950 coordinates. Only a fraction of each field was used to observe QSO candidates for this project. We list the total area corresponding to this fraction in square degrees (The full 2dF comprises $\pi$ square degrees). All 4 fields were observed for at least 6 hours, which should be sufficient to give a projected signal-to-noise of $\sim 8$ at $b_J = 22$, for seeing better than 2 arcseconds and airmass better than 1.3. The final column lists the total number of fibres placed on our QSO candidates compared to the total number of candidates in the area of observation.

QSO candidate spectra on each CCD pixel by measuring the position of salient features in our Neon lamp arc exposures in each fibre. We found that creating a composite Neon arc spectrum by sewing together longer (50s) and shorter (2s) exposures allowed us to find identifiable lines across a good range of wavelegth ($3000 - 8000$Å). An analysis of the position of sky-lines in our wavelength-calibrated spectra suggested that the major sky-lines at 5577Å and 6300Å were always within 2Å of their expected position (although it is harder to check the wavelength calibration at, say, $4000 - 5000$Å). Once all of our fields had been unpacked, clipped and flat-fielded to correct for relative throughput, sky subtraction was carried out by taking the median counts from the dedicated sky fibres in a particular field and removing this average sky spectrum from each fibre in that field. The median is used rather than the mean to account for any sky fibres that might accidentally be placed on a source. A check of the efficacy of this sky subtraction was carried out by rechecking the spectra of the dedicated sky fibres after the median sky spectra was subtracted. Comparing the sky fibres before and after subtraction suggests that typically 95 per cent of the sky is removed. Finally, each candidate QSO was checked by eye to determine what type of object it was and assigned a redshift, if extragalactic.

### 6.3.3 2dF Spectrograph Observations and Data Reduction

QSO candidates were also observed on 4 fields at the AAT over 3 dark nights, simultaneously with a second program carrying out high-resolution spectroscopy of known 2QZ QSOs. Our exposure times were fixed by this second program but fibres observing our QSO candidates were routed to a lower resolution grating (300B) to increase the signal-to-noise of our faint spectra. The 2dF exposure calculator provided on the Anglo Australian Observatory webpage suggested that a signal-to-noise ratio of $\sim 8$ required an exposure of 3 to 6 hours for seeing of 1 to 2 arcseconds for an object of magnitude $B \sim 22$. We limited our QSO candidates to the range $20.5 \leq b_J < 22$, as the intention of the second program sharing our fields was to image each of the 4 fields for at most 2 hours each night.

2dF has a number of advantages over AF2/WYFFOS for survey programs. Firstly, 400 program fibres are available and (as 2dF fibres, unlike WYFFOS fibres, can cross each other) it is usually possible to assign nearly all of the fibres to science objects (Lewis et al. 2002). Secondly, provided the relative astrometry of the sources is accurate to about 0.3 arcseconds, vignetting is not a problem for 2dF (Drinkwater et al. 2000). Note that the worst SDSS astrometry should still be better than 0.3 arcseconds (Pier et al. 2003). Finally, it is possible to configure a second field during observations, meaning that no time is lost due to reconfiguring fields during observing.

The 4 fields that were observed were not chosen specifically for our project and, consequently, some of them only partially covered the SDSS EDR (in fact, only one field completely covered the SDSS EDR region). The observing strategy adopted was to observe each of the 4 fields in turn for 2 hours at a time (in a series of 30 minute exposures), which allowed all 4 fields to be observed between evening and morning twilight (a period of 9 hours at the AAT on the nights of March 14th, 15th and 16th 2002, our observing run) together with arcs and flat field exposures. The same pattern of observations was adhered to each night, minimising the complication of fibre reconfiguration and assuring the same fibres were placed on each object each night, to yield a consistent throughput across the field. As all 4 fields for the program we shared time with were selected from the 2QZ NGC (at Declination around 0°), the airmass varied somewhat across the night. The first field observed each evening tended to be at higher airmass than the others, typically 1.3-2.0 (i.e. an altitude of 30 – 50°). Other than this first field, all of the fields were observed at low airmass ($< 1.3$).

Again, according to the constraints of the shared observational program, guide-stars were selected from sequences observed to calibrate photometry during preparation of the 2QZ input catalogue (Smith et al. 1997). About 50 fibres were reserved for sky on each field (25 of these fibres being routed to each of the two spectrographs being used). Table 6.2 summarises the observations of our QSO candidates with the 2dF spectrograph. All 3 nights were clear, with good seeing and no technical problems. Observations were carried out on our behalf by Dr. Scott Croom and Prof. Brian Boyle, of the Anglo-Australian Observatory.

The 2dF data were reduced using *2dfdr*, the dedicated 2dF data reduction pipeline (Bailey & Glazebrook 1999, Lewis et al. 2002). Most of the data reduction follows a similar methodology to our WYFFOS data reduction outlined in Section 6.3.2, above, although the advantage of dedicated reduction software is that its efficacy can constantly be rechecked at the observing site and improvements can be ongoing. Consequently, data reduction by *2dfdr* is likely superior to the WYFFOS data reduction. We assigned redshifts and identifications to our reduced spectra using *autoz*, software written by Dr. L. Miller, of Oxford University and used in the 2QZ Survey (see, e.g., Croom et al. 2001). The redshift and identification of each spectra were subsequently ratified or rejected by eye.

## 6.4 Results

### 6.4.1 Catalogues of Survey Objects

In all, 71 QSOs were identified with WYFFOS, 43 of which are definite identifications, with 28 either having low enough signal-to-noise, or enough ambiguity in the redshift determination, that the identification was in some doubt. Two of these QSOs were identified near the edge of the CCD, where the flat-fielding can be particularly bad. A total of 11 of these QSOs were identified as X-ray loud objects intended for a separate project and do not constitute part of our UVX sample. Thus the total UVX QSO sample identified with WYFFOS comprises 58 QSOs. Many additional objects were identified, including 15 Narrow Emission Line Galaxies, spanning redshifts from $z = 0.01$ to $z = 0.42$, about 25 stars or spectra with continuum emission but no discernible features, and 2 Seyfert galaxies, 2 ellipticals and a spiral in the X-ray loud sample. Tables 6.3, 6.4 and 6.5 catalogue the extragalactic objects identified with WYFFOS. All of the QSOs listed have no previous identification known to us and certainly do not appear in the most recently

| Field Centre: $17^h19^m+61°00'$, All QSOs | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (J2000) | | $\alpha$ | | | $\delta$ | | | | |
| Fibre | $h$ | $m$ | $s$ | ° | ′ | ″ | $b_J$ | $z$ | $A_{b_J}$ |
| 95 | 17 | 19 | 58.89 | +60 | 52 | 11.5 | 20.50 | 1.89 | 0.10 |
| 36 | 17 | 18 | 28.53 | +60 | 59 | 55.4 | 20.63 | 1.44? | 0.10 |
| 24 | 17 | 18 | 29.54 | +61 | 02 | 46.5 | 20.70 | 2.06 | 0.10 |
| 111 | 17 | 20 | 19.03 | +60 | 59 | 5.9 | 20.94 | 1.09 | 0.09 |
| 73 | 17 | 18 | 56.05 | +60 | 53 | 8.3 | 21.18 | 1.52 | 0.11 |
| 21 | 17 | 17 | 18.98 | +61 | 10 | 10.6 | 21.33 | 1.69 | 0.10 |
| 40 | 17 | 17 | 15.04 | +60 | 59 | 12.8 | 21.50 | 0.71? | 0.10 |
| 131v | 17 | 22 | 7.50 | +61 | 18 | 43.3 | 21.69 | 2.30? | 0.10 |
| 105 | 17 | 20 | 44.11 | +60 | 53 | 46.1 | 21.71 | 1.03? | 0.10 |
| 121 | 17 | 20 | 37.86 | +61 | 04 | 24.6 | 21.75 | 1.05 | 0.08 |
| 89 | 17 | 19 | 30.72 | +60 | 54 | 20.9 | 21.78 | 1.07 | 0.10 |
| 146 | 17 | 19 | 36.70 | +61 | 17 | 6.4 | 21.90 | 1.56 | 0.10 |
| Field Centre: $17^h25^m+59°00'$, All QSOs | | | | | | | | | |
| 146 | 17 | 25 | 21.48 | +59 | 11 | 1.5 | 20.55 | 1.95 | 0.12 |
| 114 | 17 | 26 | 40.29 | +59 | 00 | 29.7 | 20.70 | 2.46 | 0.12 |
| 81 | 17 | 25 | 24.93 | +58 | 44 | 55.6 | 20.78 | 0.66? | 0.13 |
| 90 | 17 | 25 | 44.07 | +58 | 51 | 29.6 | 21.34 | 0.93? | 0.13 |
| Field Centre: $13^h44^m-00°15'$, All QSOs | | | | | | | | | |
| 78v | 13 | 42 | 38.01 | -00 | 09 | 27.7 | 21.16 | 1.74? | 0.11 |
| 99v | 13 | 42 | 31.50 | -00 | 31 | 29.0 | 21.24 | 0.71? | 0.11 |
| 58v | 13 | 43 | 9.27 | +00 | 10 | 1.3 | 21.24 | 2.06? | 0.10 |
| 113 | 13 | 43 | 51.24 | -00 | 18 | 10.1 | 21.49 | 2.03 | 0.10 |
| 20 | 13 | 44 | 37.06 | -00 | 10 | 30.9 | 21.67 | 2.09? | 0.10 |
| 7 | 13 | 45 | 13.24 | -00 | 15 | 41.2 | 21.74 | 2.60? | 0.09 |
| 139 | 13 | 44 | 7.12 | -00 | 17 | 51.8 | 21.99 | 2.71 | 0.11 |
| 74v | 13 | 42 | 41.31 | -00 | 06 | 25.6 | 22.00 | 1.16? | 0.11 |
| 13 | 13 | 44 | 54.45 | -00 | 12 | 54.0 | 22.01 | 2.19 | 0.10 |
| 66v | 13 | 42 | 35.70 | +00 | 04 | 36.8 | 22.03 | 2.23 | 0.11 |
| 87 | 13 | 42 | 56.11 | -00 | 16 | 22.4 | 22.26 | 1.51? | 0.10 |
| 101 | 13 | 43 | 27.27 | -00 | 21 | 46.5 | 22.34 | 1.96 | 0.10 |

Table 6.3: QSOs identified (in the magnitude range $20.5 \leq b_J < 22.5$) in several fields with AF2/WYFFOS on the WHT. Field centres are listed above each table. Fields are 1° in diameter. Right Ascension ($\alpha$) and Declination ($\delta$) are J2000. The fibre number refers to the fibre the object was observed with in its field. Fibres marked "v" lay beyond 20 arcminutes of the field centre (i.e. beyond the unvignetted field of view of WYFFOS). Redshifts ($z$) marked with a question-mark are uncertain, as they are based on a single emission line in the spectrum. No listed redshift implies a missing or an unexpected emission line in the spectrum. $A_{b_J}$ denotes the predicted obscuration of the QSO due to absorption by dust in our Galaxy (Schlegel, Finkbeiner & Davis 1998).

| Field Centre: $12^h30^m+00°00'$, All QSOs | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (J2000) | $\alpha$ | | | $\delta$ | | | | | |
| Fibre | $h$ | $m$ | $s$ | ° | ′ | ″ | $b_J$ | $z$ | $A_{b_J}$ |
| 14! | 12 | 29 | 6.53 | -00 | 03 | 28.1 | 20.58 | 2.09 | 0.09 |
| 30 | 12 | 29 | 8.95 | -00 | 15 | 17.5 | 20.73 | 0.55? | 0.09 |
| 138 | 12 | 29 | 29.08 | +00 | 08 | 0.3 | 20.77 | 1.59 | 0.09 |
| 42 | 12 | 29 | 47.81 | -00 | 12 | 2.5 | 20.81 | 0.66 | 0.09 |
| 97! | 12 | 31 | 2.02 | +00 | 09 | 1.5 | 20.84 | 0.97? | 0.10 |
| 67 | 12 | 30 | 38.61 | -00 | 08 | 59.5 | 21.68 | 3.25? | 0.10 |
| 17 | 12 | 28 | 59.24 | -00 | 05 | 14.9 | 21.69 | ? | 0.09 |
| 148 | 12 | 29 | 31.66 | +00 | 04 | 8.2 | 21.71 | 2.35 | 0.09 |
| 28 | 12 | 29 | 12.74 | -00 | 12 | 37.8 | 21.75 | 1.41? | 0.09 |
| 21 | 12 | 29 | 0.19 | -00 | 07 | 35.5 | 21.97 | 2.84 | 0.09 |
| Field Centre: $12^h50^m+00°00'$, All QSOs | | | | | | | | | |
| 100v | 12 | 51 | 13.49 | -00 | 12 | 4.9 | 20.56 | 2.28 | 0.10 |
| 111 | 12 | 50 | 35.60 | -00 | 01 | 0.7 | 21.22 | 0.96? | 0.10 |
| 63 | 12 | 49 | 46.99 | -00 | 06 | 16.1 | 21.48 | 2.36 | 0.09 |
| 43 | 12 | 49 | 24.67 | -00 | 01 | 58.1 | 21.63 | 1.17? | 0.09 |
| 61 | 12 | 49 | 30.60 | -00 | 09 | 47.7 | 21.63 | ? | 0.09 |
| 115 | 12 | 51 | 17.33 | +00 | 00 | 33.7 | 21.72 | 1.56 | 0.10 |
| 121 | 12 | 50 | 53.19 | +00 | 03 | 44.9 | 21.82 | 2.45 | 0.10 |
| 82 | 12 | 50 | 14.18 | -00 | 12 | 32.5 | 21.86 | 1.23 | 0.09 |
| 130 | 12 | 50 | 43.02 | +00 | 07 | 57.1 | 21.99 | 1.23? | 0.10 |
| 88v | 12 | 50 | 45.19 | -00 | 21 | 24.3 | 22.02 | 2.18 | 0.10 |
| 38 | 12 | 49 | 50.32 | -00 | 00 | 28.6 | 22.06 | 1.05 | 0.09 |
| 59 | 12 | 49 | 57.71 | -00 | 00 | 52.8 | 22.45 | 1.35? | 0.09 |
| Field Centre: $13^h20^m+00°15'$, All QSOs | | | | | | | | | |
| 68v | 13 | 19 | 20.65 | -00 | 11 | 34.9 | 20.67 | 2.70 | 0.10 |
| 48 | 13 | 19 | 5.03 | +00 | 08 | 29.7 | 21.83 | 1.96 | 0.10 |
| Field Centre: $17^h31^m+57°00'$, All QSOs | | | | | | | | | |
| 52 | 17 | 29 | 19.65 | +57 | 00 | 19.3 | 21.08 | 1.56 | 0.15 |
| 148* | 17 | 32 | 23.68 | +57 | 10 | 36.8 | 21.16 | 2.71 | 0.18 |
| 2 | 17 | 31 | 21.42 | +57 | 05 | 54.9 | 21.21 | 1.94 | 0.17 |
| 15 | 17 | 31 | 10.86 | +57 | 06 | 31.5 | 21.48 | 1.37 | 0.17 |
| 149* | 17 | 32 | 2.20 | +57 | 10 | 3.81 | 21.67 | 1.37 | 0.18 |
| 61 | 17 | 29 | 49.92 | +56 | 57 | 3.6 | 21.82 | 0.77 | 0.15 |
| 118 | 17 | 31 | 40.40 | +56 | 57 | 44.0 | 21.88 | 1.09? | 0.15 |
| 58 | 17 | 29 | 10.16 | +56 | 57 | 7.9 | 22.33 | 2.01 | 0.15 |

Table 6.4: Formatted as for Table 6.3, for several more fields observed with AF2/WYFFOS on the WHT. Objects marked ! were observed in the 2QZ survey but were not identified, having low signal-to-noise. Objects marked * were observed at the very edge of the CCD and are not included in any of the analysis of this chapter, as the flat-fielding is particularly poor for these spectra.

Field Centre: $13^h44^m-00°15'$, All X-ray Loud Objects

| (J2000) | | $\alpha$ | | | $\delta$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Fibre | $h$ | $m$ | $s$ | ° | ′ | ″ | $b_J$ | $z$ | ID |
| 16 | 13 | 44 | 24.56 | -00 | 13 | 3.1 | 20.97 | 1.11 | QSO |
| 21 | 13 | 45 | 8.03 | -00 | 05 | 26.7 | 20.09 | 2.12? | QSO |
| 32 | 13 | 44 | 52.91 | +00 | 05 | 20.2 | 17.47 | 0.09 | Sey Gal |
| 33 | 13 | 44 | 20.82 | -00 | 04 | 58.9 | 21.26 | 0.33 | Ell gal |
| 36 | 13 | 44 | 25.94 | -00 | 00 | 56.2 | 18.78 | 1.09 | QSO |
| 40 | 13 | 44 | 20.89 | +00 | 02 | 27.0 | 20.29 | 1.87 | QSO |
| 46 | 13 | 44 | 1.94 | +00 | 00 | 3.5 | 20.22 | 0.33 | Spi Gal |
| 49 | 13 | 43 | 51.06 | +00 | 04 | 34.7 | 17.50 | 0.07 | Sey Gal |
| 57 | 13 | 43 | 29.21 | +00 | 01 | 32.7 | 21.63 | 2.34 | QSO |
| 60 | 13 | 43 | 27.29 | -00 | 01 | 38.5 | 19.73 | 0.25 | Ell gal |
| 67 | 13 | 43 | 13.36 | -00 | 05 | 1.3 | 20.79 | 1.43? | QSO |
| 125 | 13 | 44 | 12.88 | -00 | 30 | 5.5 | 20.63 | 1.48? | QSO |
| 129 | 13 | 44 | 20.07 | -00 | 31 | 10.6 | 20.56 | 2.02? | QSO |
| 131 | 13 | 44 | 23.95 | -00 | 28 | 46.5 | 21.51 | 2.22 | QSO |
| 133 | 13 | 44 | 37.40 | -00 | 32 | 37.4 | 22.05 | 2.52 | QSO |
| 135 | 13 | 44 | 24.54 | -00 | 24 | 12.8 | 21.19 | 1.76 | QSO |
| All Fields, All Galaxies | | | | | | | | | |
| 24 | 12 | 29 | 10.11 | -00 | 08 | 27.5 | 21.99 | 0.19 | NELG |
| 47 | 12 | 30 | 4.26 | -00 | 17 | 49.0 | 21.44 | 0.14 | NELG |
| 91 | 12 | 30 | 47.81 | +00 | 04 | 11.0 | 21.91 | ? | NELG |
| 77 | 12 | 50 | 3.01 | -00 | 29 | 12.3 | 22.14 | 0.37 | NELG |
| 80 | 12 | 50 | 8.23 | -00 | 17 | 35.8 | 22.43 | 0.13 | NELG |
| 127 | 12 | 51 | 35.42 | +00 | 13 | 46.7 | 21.99 | 0.24 | NELG |
| 69 | 13 | 42 | 49.57 | -00 | 01 | 36.6 | 22.15 | 0.42? | NELG |
| 19 | 17 | 17 | 39.56 | +61 | 09 | 10.4 | 21.65 | 0.29 | NELG |
| 48 | 17 | 18 | 20.13 | +60 | 57 | 37.9 | 21.91 | 0.01 | NELG |
| 130 | 17 | 21 | 38.45 | +61 | 14 | 50.6 | 21.14 | 0.24 | NELG |
| 31 | 17 | 22 | 4.46 | +59 | 06 | 56.9 | 21.10 | 0.03 | NELG |
| 119 | 17 | 26 | 43.32 | +59 | 03 | 13.1 | 22.10 | 0.18 | NELG |
| 67 | 17 | 29 | 30.03 | +56 | 52 | 8.7 | 22.06 | 0.29 | NELG |
| 99 | 17 | 31 | 40.69 | +56 | 41 | 54.2 | 21.58 | 0.38 | NELG |
| 116 | 17 | 32 | 9.20 | +56 | 53 | 56.1 | 21.58 | 0.42 | NELG |

Table 6.5: The upper panel lists X-ray loud objects in the AF2/WYFFOS subset of the survey outlined in this chapter. These objects were observed for a separate project and are not part of our UVX $20.5 < b_J < 22.5$ sample. The lower table lists non-QSO AGN identified in the AF2/WYFFOS subset of our survey. The 5 sections of the table represent the 5 different fields in the survey - the field centres are recorded in Table 6.4 and 6.3. The format is as for Table 6.4, with an extra column listing the object identification. Non-QSOs are classified as NELG (Narrow Emission Line Galaxy), Sey Gal (Seyfert Galaxy), Ell Gal (Elliptical Galaxy) or Spi Gal (Spiral Galaxy).

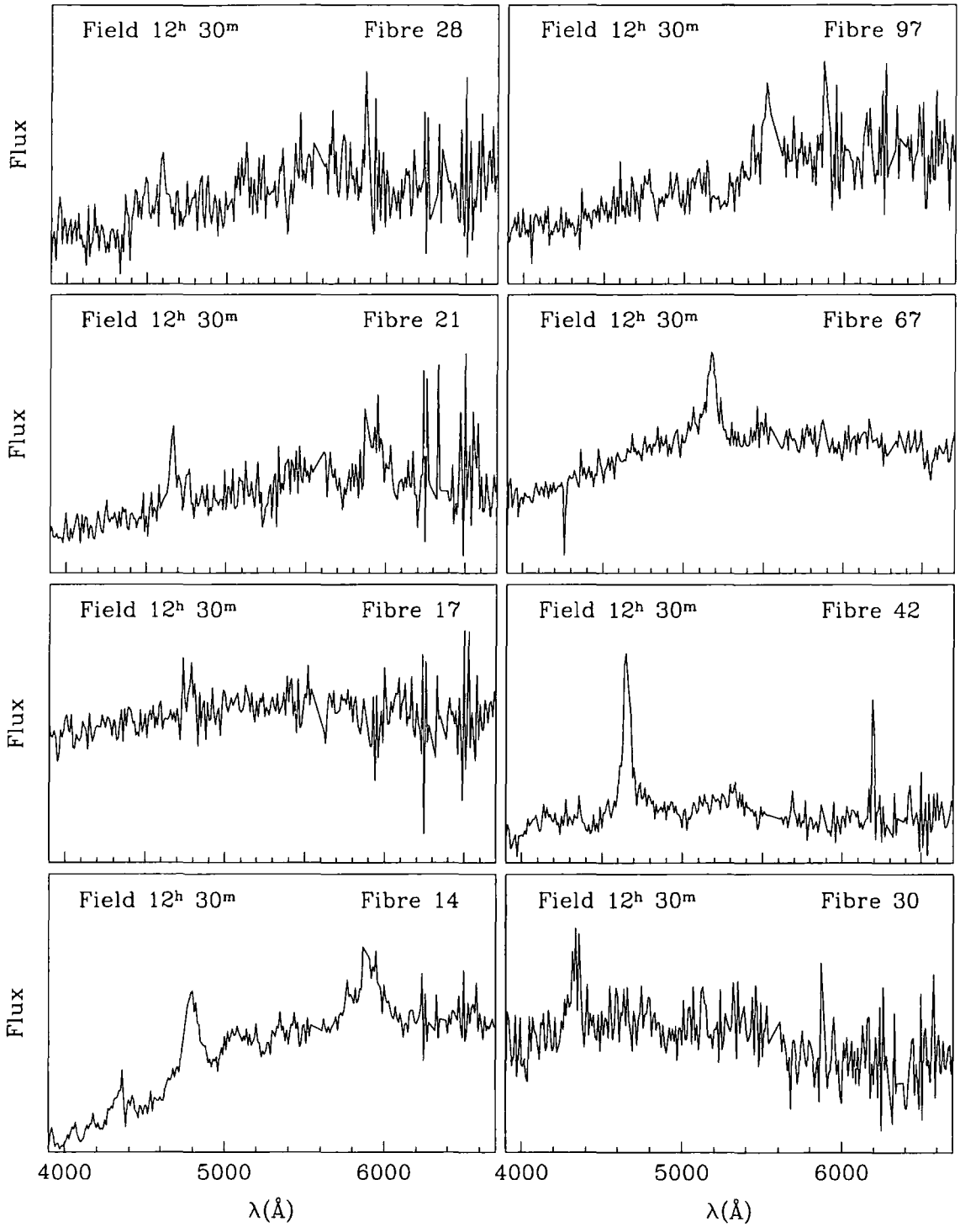


Figure 6.3: On this page, we plot QSO spectra observed with AF2/WYFFOS on the WHT. The Right Ascension of the field centre and the fibre used to observe the object are listed. The flux is in arbitrary units.

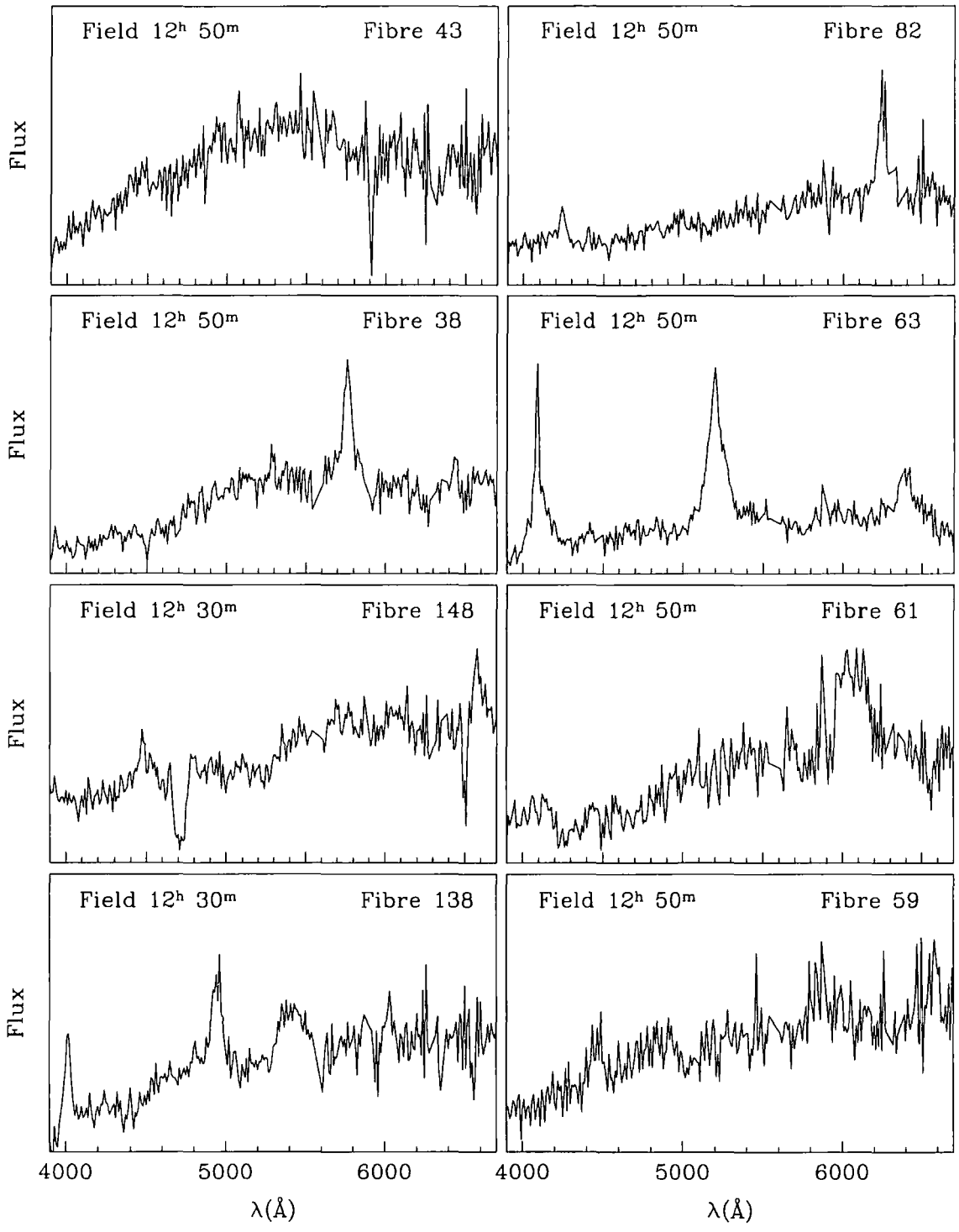


Figure 6.4: As for Fig. 6.3

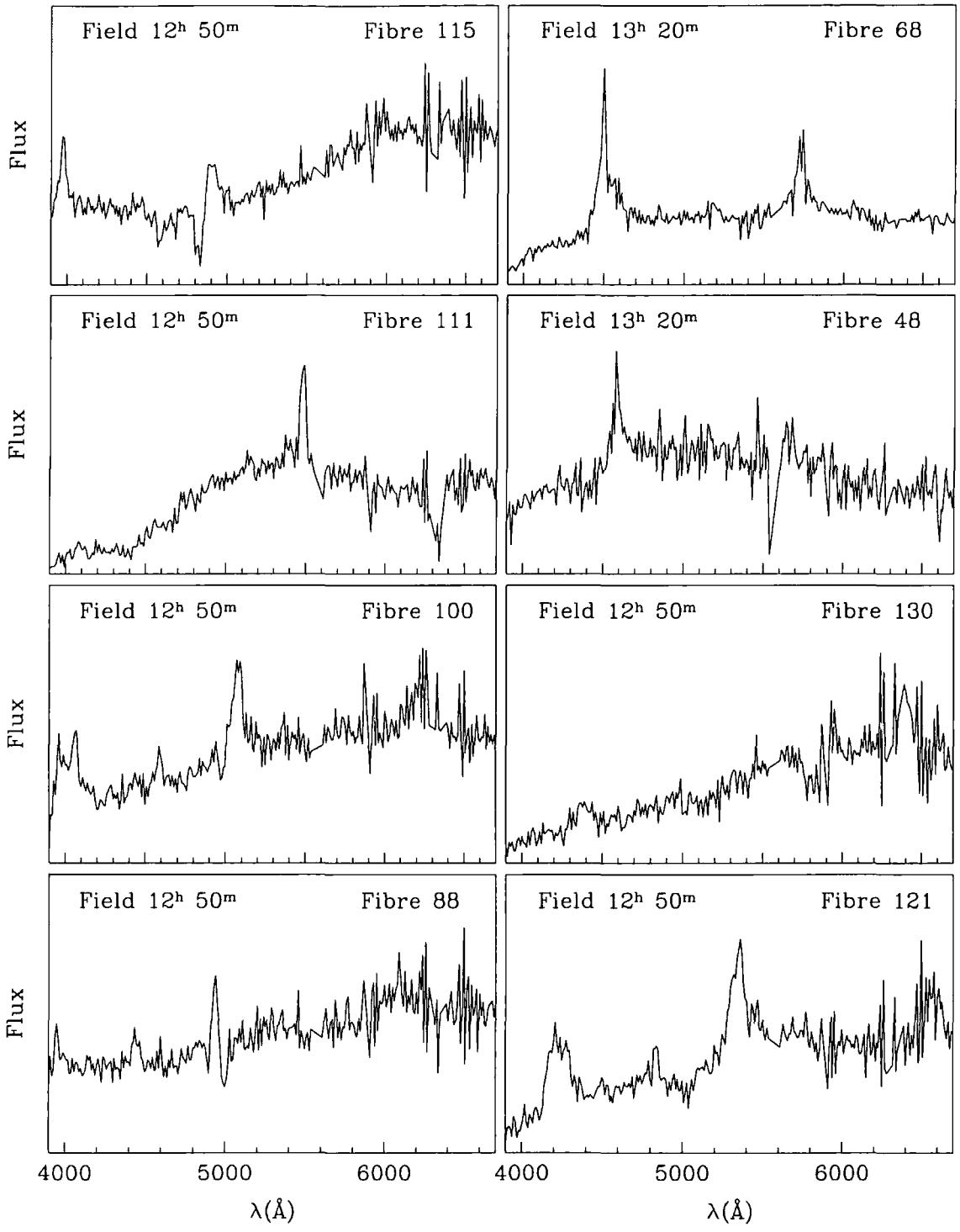


Figure 6.5: As for Fig. 6.3

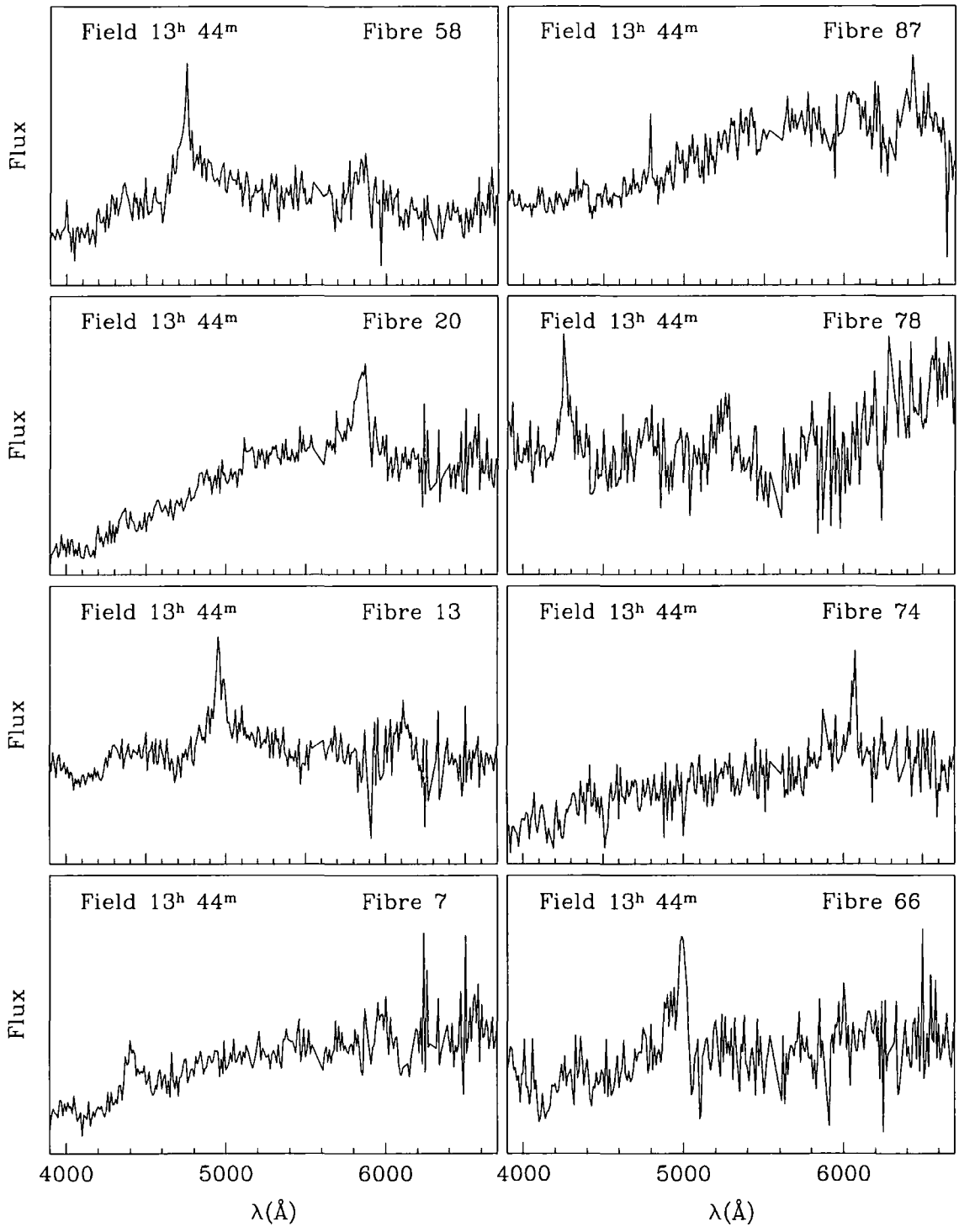


Figure 6.6: As for Fig. 6.3

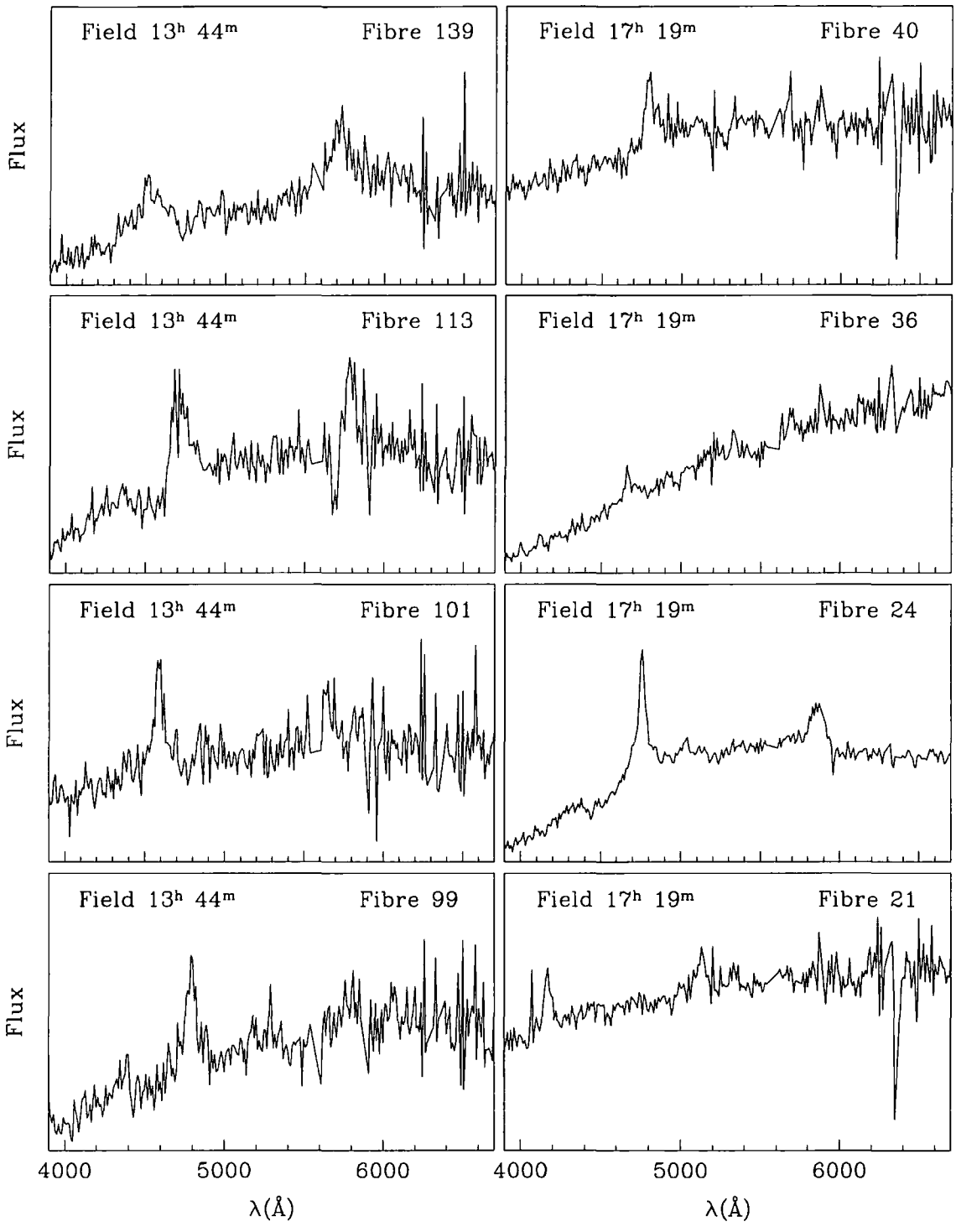


Figure 6.7: As for Fig. 6.3

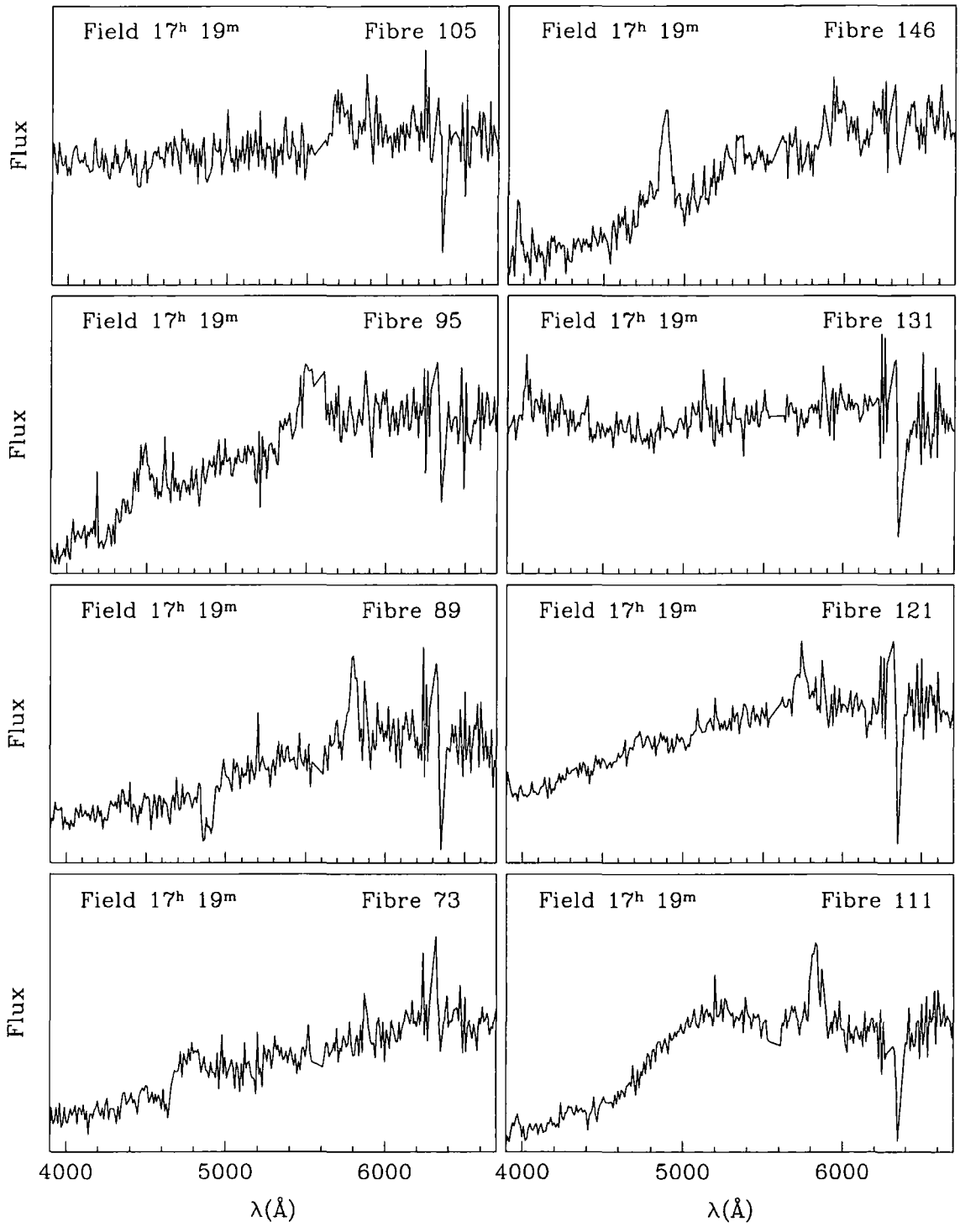


Figure 6.8: As for Fig. 6.3

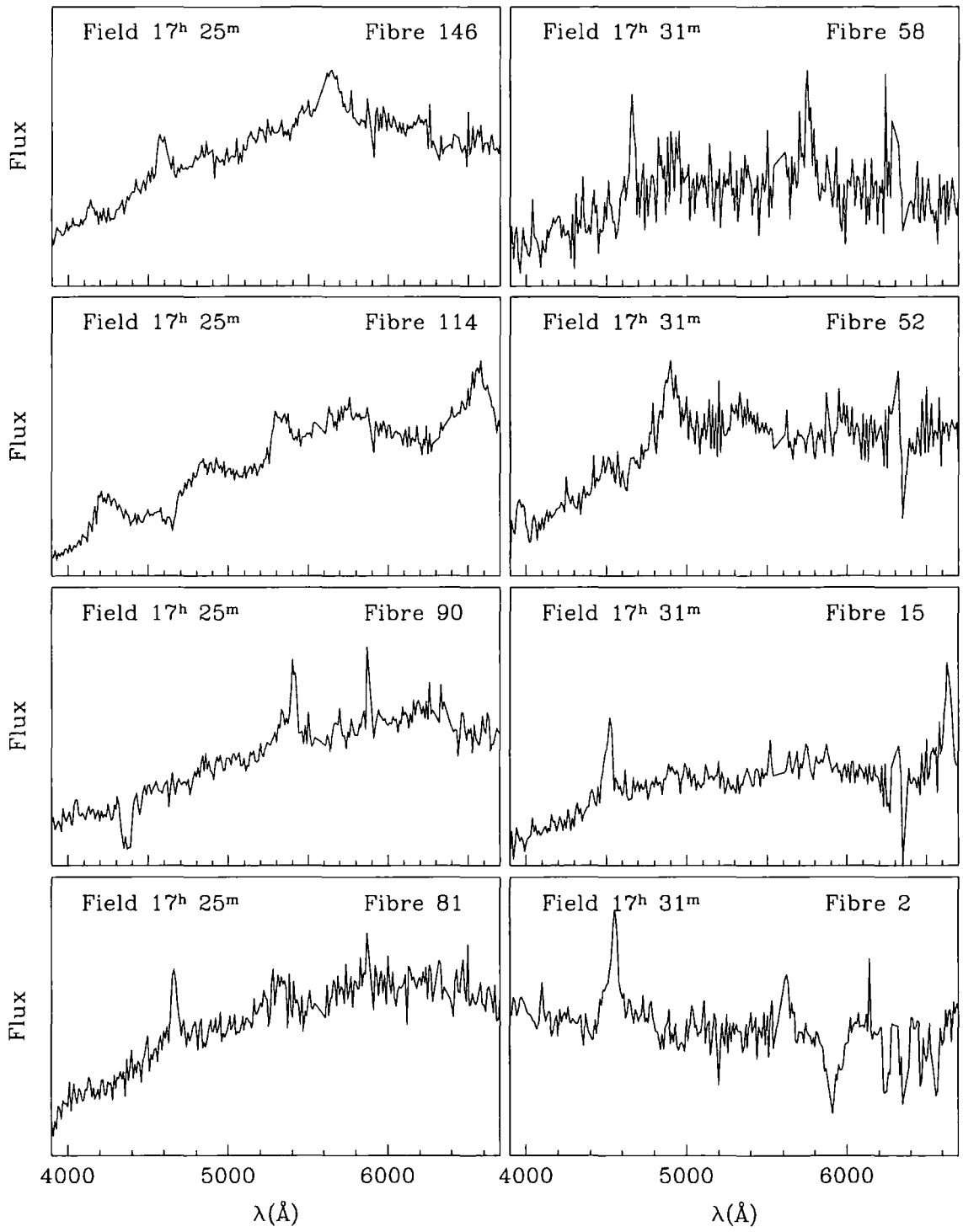


Figure 6.9: As for Fig. 6.3

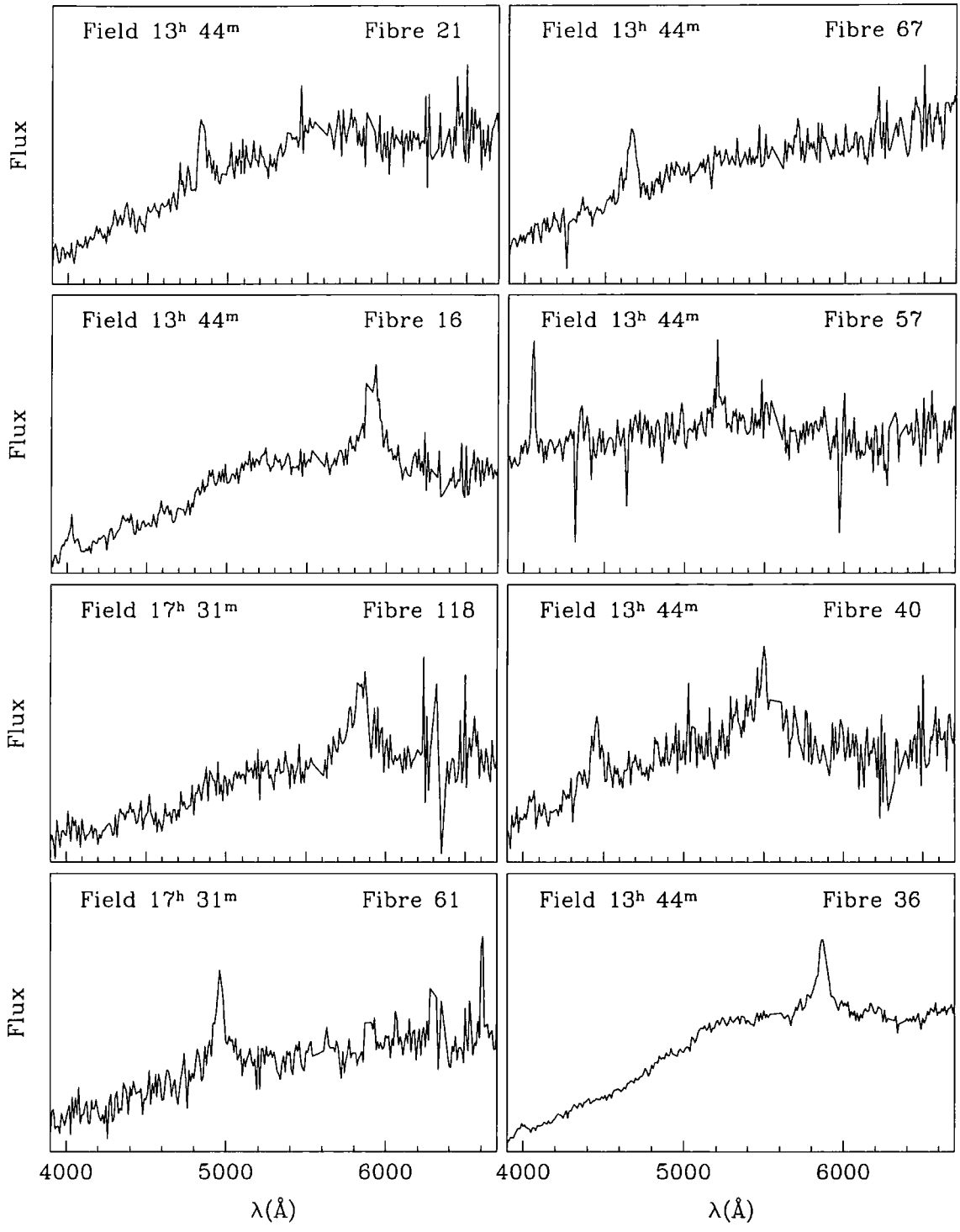


Figure 6.10: As for Fig. 6.3. The spectra for the $13^h44^m$ field are of X-ray loud QSOs. In some cases, these QSOs do not have magnitudes in the range $20.5 \leq b_J < 22.5$, and they do not constitute part of our UVX sample.

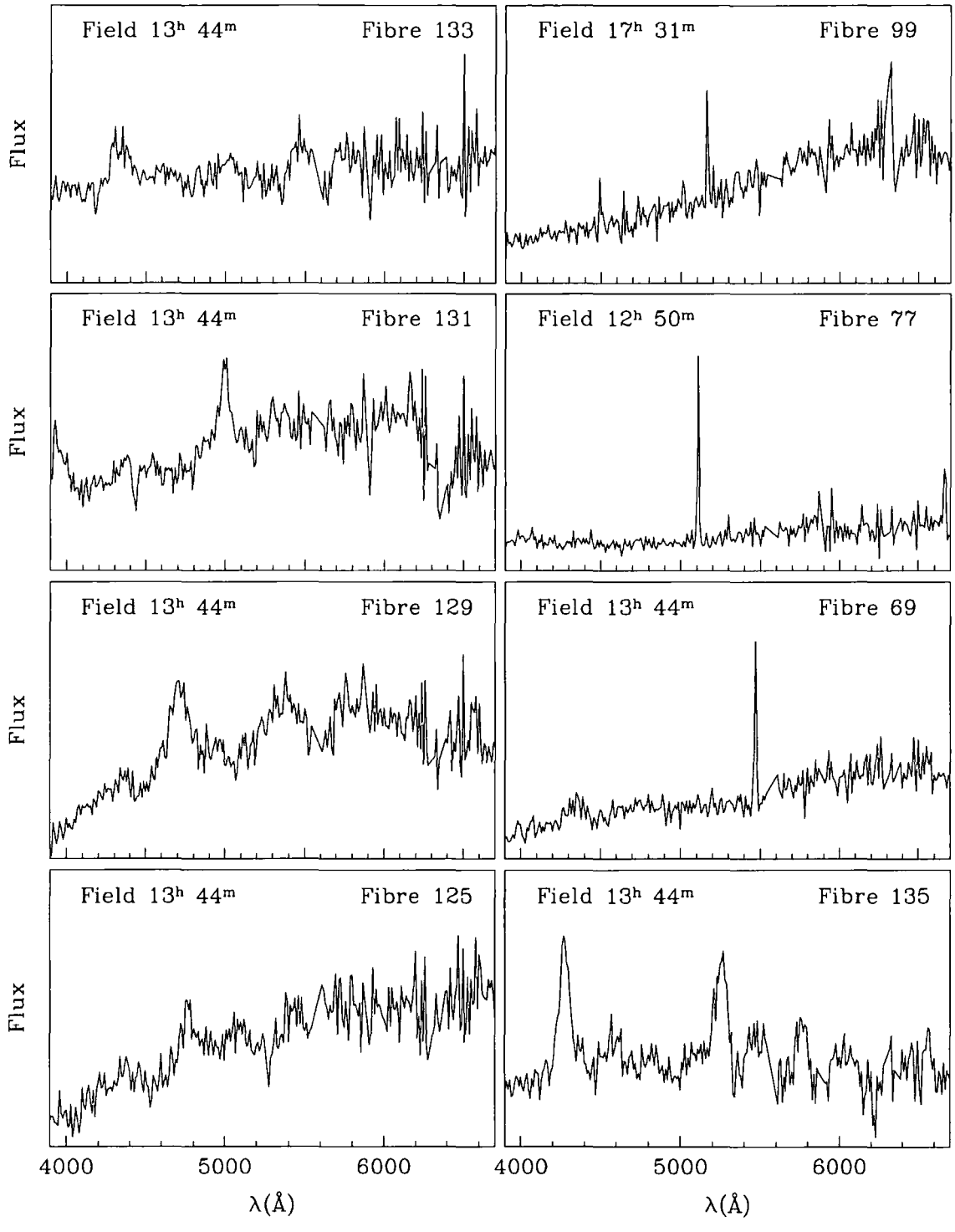


Figure 6.11: As for Fig. 6.10, except for the three upper panels in the right-hand column, which are examples of the spectra of NELGs in our survey.

updated QSO catalogue of Véron-Cetty & Véron (2001). We have estimated the error on our $b_J$ magnitudes in Section 6.2 as $\pm 0.08$. The error on our redshifts is probably no worse than $\pm 0.03$, based on the mean difference arising from measuring a QSO redshift from opposite sides of each emission line. This, of course, assumes that the position of the QSO absorption lines are of cosmological origin and not due to actual line-of-sight velocity. Obviously NELGs, having much narrower emission lines, and QSOs with more than 2 emission features or a narrow feature (such as [OIII]), will have more accurate redshifts.

In Figures 6.3 through 6.11, we display most of our reduced QSO spectra (we do not display the two QSOs that were identified near the edge of the CCD, where the flat-fielding was particularly bad). The spectra are not absolutely flux calibrated, and so the ordinate of any flux measurement is arbitrary. Some spectra in Figures 6.10 and 6.11 display the X-ray loud sample included on the $13^h44^m$ field and are not part of the main UVX QSO sample. Some of the panels in 6.11 display the typical spectra of NELGs identified in our UVX sample. All of the spectra have had the brightest sky-lines interpolated across before being displayed, occasionally, this has removed some structure from the spectra around 6300Å. One QSO identified with WYFFOS appears to have a Broad Absorption Line (BAL), the QSO identified with fibre 148 on the $12^h30^m$ field. However, many of the WYFFOS spectra observed with fibres in the range 148-150 or 1-3 show broad absorption features on all of the fields and it is extremely likely that these features arise from flat fielding problems caused, perhaps, by light loss in specific fibres that are fed to the very edges of the CCD. Only 2.4 per cent of QSOs in the 2QZ are BAL QSOs, so we would only expect $0-3$ BAL QSOs in our WYFFOS sample.

In Tables 6.6, 6.7, 6.8 and 6.9, we display the identifications and redshifts for the AGN identified with the 2dF spectrograph. We have identified 100 new QSOs with 2dF, and confirmed the identification and redshift of a further 3 QSOs from the sample of Hall et al. (1996). The 3 reidentified QSOs have redshifts that differ at most 3 per cent between our survey and that of Hall et al. (1996), consistent with our redshift error determination for the WYFFOS data. There are 4 BAL QSOs in our sample, consistent with the expected number from the 2QZ survey. We additionally identify 19 Narrow Emission Line Galaxies (note that some of these are actually specified as Low Ionisation Nuclear Emission Regions by *autoz*) and around 80 stars or objects with continuum emission but no obvious features. For brevity, we don't display the spectra of all of these objects but they are available on request.

Field Centre: $14^h00^m35.5^s-00°55'29''$, All QSOs

| (B1950) | $\alpha$ | | | $\delta$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Fibre | $h$ | $m$ | $s$ | ° | ′ | ″ | $b_J$ | $z$ | $A_{b_J}$ |
| 91! | 13 | 58 | 17.39 | -00 | 47 | 42.9 | 20.65 | 2.04 | 0.18 |
| 68 | 13 | 59 | 51.99 | -00 | 31 | 31.1 | 20.76 | 1.55 | 0.19 |
| 53 | 14 | 00 | 2.48 | -00 | 00 | 13.3 | 20.78 | 2.34 | 0.20 |
| 41 | 14 | 00 | 56.51 | -00 | 21 | 19.4 | 20.80 | 2.73 | 0.20 |
| 20 | 14 | 02 | 29.18 | -00 | 42 | 59.2 | 20.86 | 2.14 | 0.22 |
| 49 | 14 | 01 | 28.23 | -00 | 03 | 14.6 | 20.92 | 2.40 | 0.21 |
| 112* | 13 | 59 | 14.65 | -00 | 49 | 9.8 | 20.97 | 2.10 | 0.19 |
| 193 | 14 | 01 | 50.84 | -00 | 48 | 16.9 | 20.98 | 1.65 | 0.22 |
| 113 | 13 | 59 | 22.18 | -00 | 45 | 12.9 | 21.12 | 1.17 | 0.18 |
| 167 | 14 | 01 | 0.80 | -00 | 52 | 13.2 | 21.17 | 2.15 | 0.20 |
| 51 | 14 | 00 | 33.16 | -00 | 11 | 3.2 | 21.19 | 1.55 | 0.21 |
| 65 | 13 | 59 | 53.41 | -00 | 19 | 35.9 | 21.34 | 1.61 | 0.20 |
| 77 | 13 | 58 | 4.31 | -00 | 18 | 26.7 | 21.36 | 1.34 | 0.19 |
| 84 | 13 | 57 | 37.32 | -00 | 37 | 43.6 | 21.38 | 1.07 | 0.18 |
| 71* | 13 | 59 | 54.46 | -00 | 26 | 7.1 | 21.44 | 1.98 | 0.19 |
| 39 | 14 | 01 | 57.83 | -00 | 11 | 8.0 | 21.44 | 1.54 | 0.21 |
| 198 | 14 | 02 | 42.69 | -00 | 49 | 8.9 | 21.57 | 0.95 | 0.25 |
| 38* | 14 | 01 | 55.43 | -00 | 03 | 49.1 | 21.58 | 2.21 | 0.20 |
| 27 | 14 | 01 | 41.98 | -00 | 23 | 48.3 | 21.64 | 1.20 | 0.20 |
| 103! | 13 | 58 | 47.26 | -00 | 50 | 23.5 | 21.71 | 1.35 | 0.19 |
| 54 | 14 | 00 | 24.41 | -00 | 09 | 23.7 | 21.80 | 1.20 | 0.21 |
| 48 | 14 | 01 | 23.96 | -00 | 15 | 23.0 | 21.81 | 1.55 | 0.21 |
| 93! | 13 | 58 | 9.78 | -00 | 39 | 41.8 | 21.81 | 1.38 | 0.18 |
| 46 | 14 | 01 | 13.29 | -00 | 16 | 42.6 | 21.84 | 1.41 | 0.20 |
| 110 | 13 | 59 | 38.05 | -00 | 51 | 44.4 | 21.88 | 2.35 | 0.18 |
| 57 | 14 | 00 | 26.45 | -00 | 22 | 18.0 | 21.88 | 1.36 | 0.20 |
| 50 | 14 | 01 | 19.02 | -00 | 11 | 10.8 | 22.00 | 1.67 | 0.21 |

Table 6.6: QSOs identified (in the magnitude range $20.5 \leq b_J < 22$) in the 2dF field centred at $14^h00^m35.5^s-00°55'29''$. The diameter of the field is 2°. Right Ascension ($\alpha$) and Declination ($\delta$) in this table are in B1950 coordinates. The fibre number refers to the 2dF fibre the object was observed with. Fibre numbers marked with a * are Broad Absorption Line QSOs. Redshifts ($z$) marked with a question-mark are uncertain, as they are based on a low signal-to-noise spectra, or on a single emission line in the spectrum. $A_{b_J}$ denotes the expected obscuration of the listed QSO due to absorption by dust in our Galaxy (Schlegel, Finkbeiner & Davis 1998). The objects marked ! were previously identified in the sample of Hall et al. (1996), all with redshifts measured within 3 per cent of our own.

Field Centre: $12^h54^m18.1^s{+}00°03'09''$, all QSOs

| (B1950) | $\alpha$ | | | $\delta$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Fibre | $h$ | $m$ | $s$ | ° | $'$ | $''$ | $b_J$ | $z$ | $A_{b_J}$ |
| 83 | 12 | 51 | 58.38 | +00 | 25 | 40.0 | 20.66 | 1.38 | 0.09 |
| 171 | 12 | 56 | 19.80 | -00 | 31 | 37.2 | 20.86 | 1.79 | 0.08 |
| 68 | 12 | 53 | 56.44 | +00 | 29 | 20.3 | 20.92 | 1.29 | 0.08 |
| 143 | 12 | 53 | 2.10 | -00 | 07 | 46.2 | 20.94 | 1.28 | 0.10 |
| 182 | 12 | 56 | 59.48 | -00 | 13 | 7.9 | 20.98 | 1.52 | 0.10 |
| 155 | 12 | 55 | 5.81 | -00 | 43 | 10.8 | 21.02 | 0.66 | 0.09 |
| 59 | 12 | 54 | 11.85 | +00 | 38 | 16.0 | 21.11 | 1.99 | 0.08 |
| 71 | 12 | 53 | 8.28 | +00 | 26 | 58.2 | 21.17 | 1.27 | 0.08 |
| 144 | 12 | 53 | 3.21 | +00 | 00 | 19.4 | 21.18 | 1.46 | 0.10 |
| 16 | 12 | 57 | 6.68 | +00 | 23 | 15.3 | 21.18 | 0.63? | 0.10 |
| 76 | 12 | 52 | 22.66 | +00 | 28 | 34.5 | 21.33 | 2.30 | 0.09 |
| 89 | 12 | 52 | 52.31 | +00 | 20 | 32.3 | 21.36 | 1.52 | 0.09 |
| 173 | 12 | 56 | 11.98 | -00 | 32 | 5.2 | 21.37 | 0.89 | 0.07 |
| 9 | 12 | 57 | 47.73 | +00 | 19 | 49.6 | 21.38 | 1.06 | 0.09 |
| 102 | 12 | 52 | 55.38 | +00 | 15 | 54.1 | 21.38 | 2.27 | 0.10 |
| 146 | 12 | 54 | 35.01 | -00 | 14 | 52.1 | 21.38 | 2.56 | 0.13 |
| 147 | 12 | 53 | 37.76 | -00 | 34 | 31.6 | 21.42 | 1.20 | 0.07 |
| 82 | 12 | 51 | 35.20 | +00 | 23 | 47.4 | 21.43 | 1.24? | 0.09 |
| 113 | 12 | 51 | 33.39 | -00 | 15 | 44.6 | 21.49 | 1.69 | 0.09 |
| 8 | 12 | 57 | 27.44 | +00 | 11 | 43.5 | 21.52 | 0.66? | 0.10 |
| 125 | 12 | 51 | 52.15 | -00 | 38 | 4.6 | 21.59 | 2.01 | 0.08 |
| 168 | 12 | 56 | 15.05 | -00 | 37 | 31.4 | 21.61 | 0.98? | 0.07 |
| 62 | 12 | 52 | 50.34 | +00 | 36 | 51.1 | 21.62 | 2.33 | 0.08 |
| 41 | 12 | 55 | 18.95 | +00 | 46 | 49.1 | 21.63 | 2.43? | 0.08 |
| 84 | 12 | 53 | 0.58 | +00 | 23 | 45.0 | 21.66 | 1.97 | 0.09 |
| 45 | 12 | 54 | 48.79 | +00 | 50 | 18.1 | 21.67 | 1.61 | 0.07 |
| 115 | 12 | 51 | 56.48 | -00 | 13 | 14.2 | 21.67 | 0.62 | 0.09 |
| 57 | 12 | 53 | 34.57 | +00 | 52 | 41.5 | 21.73 | 1.97 | 0.08 |
| 74 | 12 | 53 | 36.68 | +00 | 26 | 35.6 | 21.73 | 2.56 | 0.08 |
| 176 | 12 | 56 | 18.33 | -00 | 30 | 2.8 | 21.78 | 1.84 | 0.08 |
| 92 | 12 | 54 | 5.72 | +00 | 24 | 45.3 | 21.84 | 1.54 | 0.09 |
| 100 | 12 | 52 | 49.14 | +00 | 12 | 9.7 | 21.84 | 2.29 | 0.10 |
| 150 | 12 | 54 | 39.74 | -00 | 18 | 28.5 | 21.87 | 1.82 | 0.13 |
| 70 | 12 | 54 | 22.19 | +00 | 30 | 26.6 | 21.94 | 1.81 | 0.08 |
| 96 | 12 | 51 | 56.76 | +00 | 03 | 7.3 | 21.94 | 2.19 | 0.11 |
| 124 | 12 | 51 | 50.81 | -00 | 37 | 9.4 | 21.95 | 2.38 | 0.08 |
| 21 | 12 | 55 | 38.98 | +00 | 17 | 35.8 | 21.96 | 2.05 | 0.09 |
| 91 | 12 | 53 | 30.57 | +00 | 24 | 29.3 | 21.96 | 2.14 | 0.08 |
| 157* | 12 | 55 | 34.40 | -00 | 08 | 58.1 | 21.99 | 1.87 | 0.12 |

Table 6.7: Formatted as for Table 6.6, for the 2dF field centred at $12^h54^m18.1^s{+}00°03'09''$.

Field Centre: $10^h34^m57.4^s{+}01°23'17''$, all QSOs

| (B1950) | | $\alpha$ | | | $\delta$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Fibre | $h$ | $m$ | $s$ | ° | ′ | ″ | $b_J$ | $z$ | $A_{b_J}$ |
| 191 | 10 | 38 | 24.34 | +01 | 09 | 51.6 | 20.78 | 1.65 | 0.24 |
| 147 | 10 | 34 | 35.40 | +00 | 20 | 51.6 | 20.92 | 2.27 | 0.30 |
| 200 | 10 | 37 | 57.58 | +01 | 18 | 12.4 | 20.94 | 1.58 | 0.21 |
| 156 | 10 | 36 | 7.72 | +01 | 19 | 1.8 | 21.10 | 0.83 | 0.20 |
| 168 | 10 | 37 | 4.02 | +00 | 32 | 56.9 | 21.14 | 2.35 | 0.24 |
| 193 | 10 | 37 | 49.54 | +01 | 10 | 33.8 | 21.15 | 1.80 | 0.24 |
| 139 | 10 | 33 | 47.84 | +00 | 46 | 0.9 | 21.16 | 2.28 | 0.26 |
| 173 | 10 | 37 | 50.65 | +00 | 40 | 34.5 | 21.25 | 1.70 | 0.25 |
| 180 | 10 | 37 | 21.87 | +00 | 51 | 14.0 | 21.27 | 0.68 | 0.26 |
| 121 | 10 | 32 | 47.36 | +00 | 59 | 13.0 | 21.36 | 0.91? | 0.24 |
| 171 | 10 | 36 | 26.32 | +00 | 53 | 56.5 | 21.46 | 1.41 | 0.26 |
| 178 | 10 | 37 | 50.49 | +00 | 52 | 5.8 | 21.50 | 2.47 | 0.26 |
| 101 | 10 | 32 | 55.96 | +01 | 16 | 5.6 | 21.50 | 1.74 | 0.21 |
| 154 | 10 | 35 | 41.31 | +00 | 26 | 32.9 | 21.54 | 1.19 | 0.27 |
| 184 | 10 | 36 | 39.54 | +01 | 02 | 34.0 | 21.58 | 1.44 | 0.25 |
| 182 | 10 | 38 | 5.02 | +01 | 03 | 34.2 | 21.65 | 1.72 | 0.25 |
| 114 | 10 | 31 | 26.48 | +00 | 54 | 31.8 | 21.66 | 0.88 | 0.27 |
| 160 | 10 | 35 | 46.91 | +00 | 39 | 39.4 | 21.68 | 1.81 | 0.26 |
| 161 | 10 | 35 | 26.95 | +00 | 54 | 43.2 | 21.70 | 2.07 | 0.25 |
| 172 | 10 | 36 | 58.10 | +00 | 45 | 5.6 | 21.73 | 2.22 | 0.26 |
| 195 | 10 | 37 | 4.52 | +01 | 17 | 3.9 | 21.75 | 1.44 | 0.22 |
| 141 | 10 | 34 | 23.17 | +00 | 42 | 22.4 | 21.86 | 0.93 | 0.26 |
| 176 | 10 | 36 | 19.31 | +00 | 56 | 1.2 | 21.96 | 2.11 | 0.26 |

Field Centre: $11^h05^m46.7^s{-}01°26'20''$, all QSOs

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 29 | 11 | 08 | 1.04 | -00 | 49 | 0.6 | 20.51 | 1.62 | 0.14 |
| 40 | 11 | 06 | 58.89 | -00 | 48 | 43.8 | 20.64 | 0.65 | 0.13 |
| 39 | 11 | 06 | 56.79 | -00 | 46 | 59.2 | 20.98 | 1.00 | 0.13 |
| 75 | 11 | 04 | 6.09 | -00 | 53 | 18.0 | 21.00 | 1.49 | 0.18 |
| 53 | 11 | 05 | 10.57 | -00 | 27 | 35.1 | 21.06 | 0.88 | 0.18 |
| 33 | 11 | 07 | 33.45 | -00 | 42 | 17.7 | 21.22 | 1.66 | 0.15 |
| 67 | 11 | 04 | 14.22 | -00 | 50 | 43.8 | 21.24 | 1.83 | 0.18 |
| 60 | 11 | 04 | 39.12 | -00 | 42 | 12.5 | 21.33 | 2.14 | 0.19 |
| 49 | 11 | 06 | 30.49 | -00 | 49 | 14.1 | 21.51 | 2.26 | 0.14 |
| 50 | 11 | 06 | 34.43 | -00 | 38 | 43.2 | 21.72 | 0.45? | 0.14 |
| 54 | 11 | 05 | 28.03 | -00 | 37 | 15.4 | 21.75 | 1.89 | 0.17 |
| 43 | 11 | 06 | 3.64 | -00 | 30 | 4.2 | 21.79 | 0.99? | 0.16 |
| 76 | 11 | 03 | 44.26 | -00 | 51 | 49.9 | 21.82 | 0.68 | 0.19 |
| 37 | 11 | 07 | 23.89 | -00 | 32 | 0.9 | 21.83 | 1.47 | 0.16 |

Table 6.8: Formatted as for Table 6.6, for the two 2dF fields centred at $10^h34^m57.4^s{+}01°23'17''$ and $11^h05^m46.7^s{-}01°26'20''$.

All Fields, All Galaxies

| (B1950) | | $\alpha$ | | | $\delta$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Fibre | $h$ | $m$ | $s$ | ° | ′ | ″ | $b_J$ | $z$ | ID |
| 117 | 10 | 31 | 31.74 | +00 | 52 | 19.8 | 21.97 | 0.44 | NELG |
| 118 | 10 | 31 | 34.82 | +00 | 52 | 5.2 | 21.64 | 0.10 | NELG |
| 158 | 10 | 35 | 19.54 | +00 | 50 | 29.4 | 21.86 | 0.19 | NELG |
| 170 | 10 | 35 | 55.50 | +01 | 00 | 6.6 | 21.65 | 0.16 | NELG |
| 183 | 10 | 36 | 55.78 | +01 | 07 | 34.9 | 21.58 | 0.11 | NELG |
| 145 | 10 | 34 | 22.73 | +00 | 29 | 22.7 | 21.87 | 0.03 | LINER |
| 61 | 11 | 04 | 50.70 | -00 | 47 | 47.1 | 21.84 | 0.28 | NELG |
| 4 | 12 | 57 | 4.31 | +00 | 11 | 34.0 | 21.62 | 0.58 | NELG |
| 12 | 12 | 57 | 3.62 | +00 | 18 | 46.0 | 21.82 | 0.31 | NELG |
| 14 | 12 | 56 | 22.94 | +00 | 17 | 57.2 | 21.95 | 0.06 | NELG |
| 72 | 12 | 53 | 11.58 | +00 | 30 | 49.7 | 21.27 | 0.20 | NELG |
| 152 | 12 | 54 | 55.94 | -00 | 46 | 40.8 | 21.59 | 0.14 | NELG |
| 160 | 12 | 55 | 10.05 | -00 | 39 | 17.2 | 21.86 | 0.12 | NELG |
| 17 | 12 | 57 | 35.74 | +00 | 31 | 57.7 | 21.38 | 0.14? | LINER |
| 149 | 12 | 53 | 36.28 | -00 | 18 | 28.6 | 20.97 | 0.16 | LINER |
| 42 | 14 | 00 | 55.94 | -00 | 00 | 29.3 | 21.73 | 0.02 | NELG |
| 74 | 13 | 58 | 8.62 | -00 | 15 | 55.1 | 21.90 | 0.01 | LINER |
| 104 | 14 | 00 | 28.56 | -00 | 51 | 21.7 | 21.42 | 0.03 | LINER |
| 115 | 13 | 59 | 56.95 | -00 | 43 | 39.6 | 21.06 | 0.05 | LINER |

Table 6.9: Other AGN identified in the 2dF subset of the survey outlined in this chapter. The 4 sections of the table represent the 4 different 2dF fields in the survey - the field centres are noted in earlier tables. Right Ascension ($\alpha$) and Declination ($\delta$) are in B1950 coordinates. The fibre number refers to the 2dF fibre that the object was observed with in its field. Redshifts ($z$) marked with a question-mark are uncertain, usually as they are based on a single emission line in the spectrum, or because an emission line that would be expected at that redshift was not found, or an unexpected emission line threw doubt on the redshift determination. Objects are classified as either NELG (Narrow Emission Line Galaxy) or LINER (Low Ionisation Nuclear Emission Region).

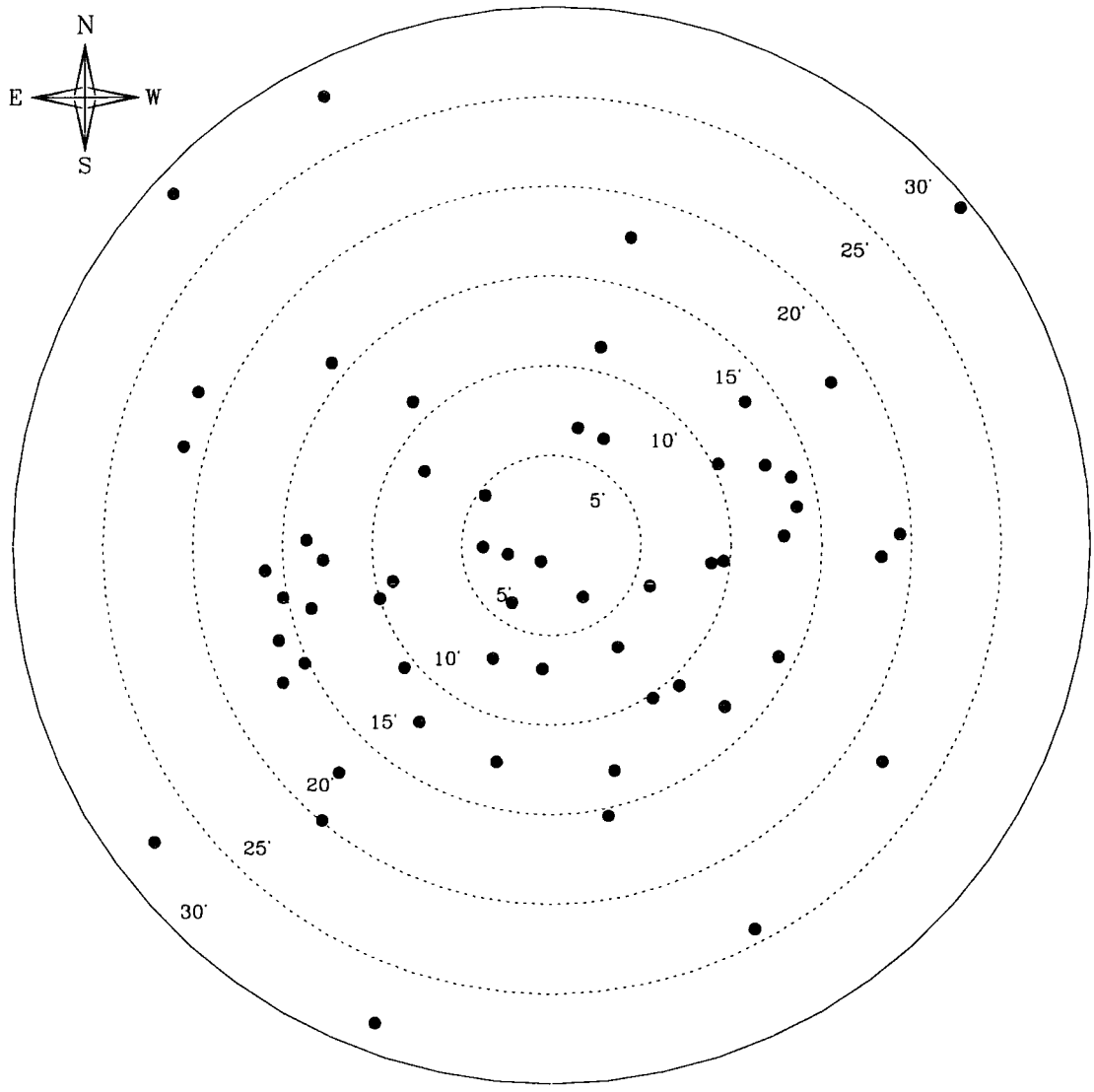


Figure 6.12: QSOs identified using WYFFOS by position on the observing plate, for the 58 UVX QSOs identifed in our sample. WYFFOS suffers some vignetting outside of 40 arcminutes diameter on the plate that can cause light loss in fibres positioned outside this radius. Note that the ratio of the total number of science fibres placed within 20 arcminutes of the field centre to the number placed across the entire field is fairly represented by the relative areas (see, e.g. Table 6.1). It is obvious that disproportionately few QSOs are observed outside the central 40 arcminutes of a WYFFOS field.

### 6.4.2 Survey Completeness and QSO Number Counts by Magnitude

We noted in Section 6.3.1 that WYFFOS can suffer considerable vignetting at a radius of more than 20 arcminutes from the field centre. In Fig 6.12, we check the extent of this effect on our QSO sample, by plotting the positon of the 58 UVX QSOs discovered with WYFFOS against the position of the fibre placed upon that QSO on the observing plate. Table 6.1 demonstrates that the relative fibre placement at 30 and 20 arcminutes from the plate centre is fairly represented by the relative areas on the plate. Blatantly, vignetting is a severe problem for this instrument (at least, observing at the faint magnitudes we're considering). In a sample of 58 QSOs, we would expect 32 QSOs to be found outside of the central 40 arcminutes of the plate and 26 QSOs within this diameter. In actuality, 49 of the QSOs found are within 20 arcminutes of the plate centre and only 9 outside this radius. Assuming Poisson statistics, the chance of this situation arising at random is less than 1 per cent. Further, now considering only the central 40 arcminutes of each WYFFOS plate, we would expect a total of 28 QSOs from our WYFFOS sample to be observed in the central 30 arcminutes diameter of the plate and 21 to lie within a radius of 15-20 arcminutes from the field centre, however, in actuality 36 QSOs lie within 15 arcminutes of the field centre, with only 13 at a radius of 15-20 arcminutes. This suggests there is still some incompleteness due to vignetting outside a diameter of 30 arcminutes on the plate but only at small significance. For these reasons, we shall only analyse the statistics of objects that lie within 20 arcminutes of the centre of a WYFFOS field. In Section 6.4.1 we noted that we had discarded two objects that lay near the edge of the CCD, where the flat-fielding was particularly bad. These objects were both observed further than 20 arcminutes from the WYFFOS plate centre, so will not influence our statistical analysis. A similar check of radial completeness on the $12^h54^m$ 2dF field (which is the only field observed with 2dF that lies entirely within the SDSS EDR) suggests no obvious radial dependence of survey completeness, although, other authors have found there may be a small variation in completeness across 2dF plates, particularly at the eastern edge of fields, from which fibres are fed to the edges of the CCD (Croom et al. 2001, Lewis et al. 2002).

Our QSO candidates were targeted using the same UVX method as the 2QZ, which is usually assumed to be about 65-90 per cent complete across the redshift range $0.3 < z < 2.2$ and to fairly sample QSOs in this redshift range (Véron 1983, Graham, Clowes & Campusano 1999, Rengstorf et al. 2001). Although it isn't simple to directly determine

| Field ($\alpha$ B1950) | | Area | OC* | IDs binned by $b_J$ as QSO/non-QSO/no-ID | | | |
|---|---|---|---|---|---|---|---|
| $h$ | $m$ | *sq.deg.* | % | 20.75 | 21.25 | 21.75 | Total |
| 10 | 34 | 1.55 | 37.6 | 7/1/4 | 13/4/12 | 3/5/16 | 23/16/32 |
| 11 | 05 | 0.56 | 37.4 | 4/3/2 | 4/2/7 | 5/1/8 | 14/6/17 |
| 12 | 54 | 3.14 | 24.2 | 6/13/6 | 14/18/20 | 19/2/27 | 39/54/53 |
| 14 | 00 | 1.48 | 30.8 | 10/6/6 | 8/12/13 | 8/5/13 | 27/25/32 |
| Total | | 6.73 | 29.0 | 27/23/18 | 39/40/52 | 35/32/64 | 103/101/134 |
| 10 | 34 | 1.55 | 37.6 | (7/1/4) | (12/5/16) | (3/3/18) | (22/9/40) |
| 11 | 05 | 0.56 | 37.4 | (4/0/5) | (4/1/8) | (3/1/10) | (12/2/23) |
| 12 | 54 | 3.14 | 24.2 | (6/3/16) | (11/5/36) | (17/5/45) | (34/14/98) |
| 14 | 00 | 1.48 | 30.8 | (10/1/11) | (8/2/23) | (8/2/16) | (27/6/51) |
| Total | | 6.73 | 29.0 | (27/5/36) | (35/13/83) | (31/11/89) | (95/31/212) |
| $0.3 < z < 2.2$ | | 6.73 | 29.0 | 21(21) | 32(28) | 29(26) | 84(77) |

Table 6.10: Numbers of objects observed in the 2dF part of the survey described in this chapter by their eventual identification. We list the fields by the Right Ascension, $\alpha$ of the field centre (B1950 coordinates) and provide the area of each field covered by the SDSS EDR, from which our QSO candidates were selected. The numbers of objects are listed in bins of 0.5 magnitude ($b_J$) denoted by the bin centre. Magnitudes of objects have been corrected for dust before producing this table using the maps of Schlegel, Finkbeiner & Davis (1998). We list the objects as Number of QSOs identified/Number of non-QSOs identifed/Number of objects without an identification. Numbers in brackets are the numbers of objects that have absolutely certain identifications (i.e. those not listed with a question-mark in previous tables). The total numbers of objects may exceed the sum of the numbers in the columns as a few objects have magnitudes brighter than considered by this table. The spectroscopic completeness of any field in any bin may be calculated as the total number of objects identified divided by the total number of objects that were observed with a fibre.*"OC" denotes observational completeness - the fraction of the total number of QSO candidates in a field that were observed with a fibre. The final row lists the number of QSOs that have a redshift in the range 0.3 to 2.2 (the redshift range for which the UVX selection method is designed to fairly sample QSOs).

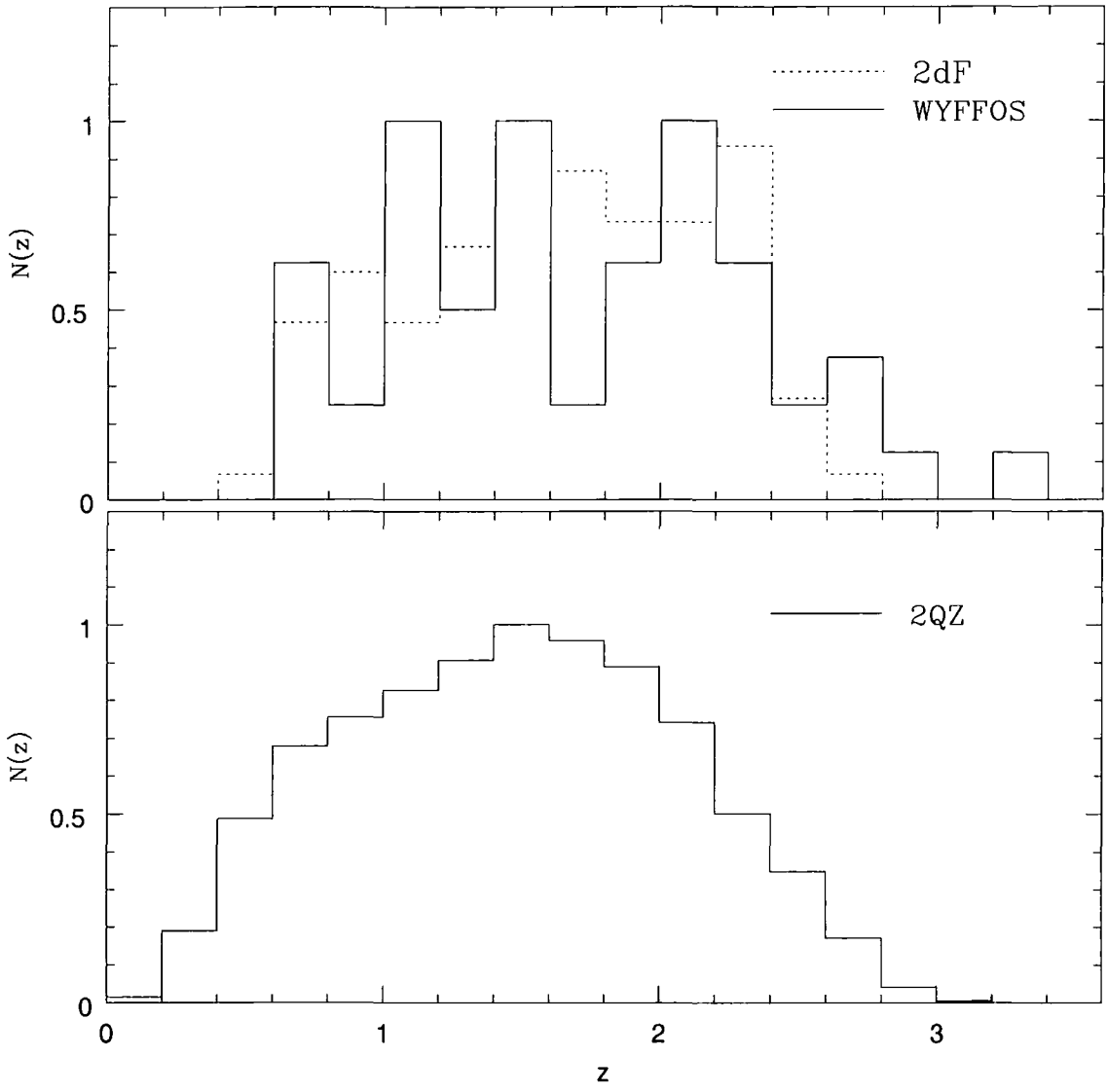


Figure 6.13: QSOs identified using both WYFFOS and the 2dF by redshift. The upper panel shows our survey data. The lower panel shows the redshift distribution of the 2dF QSO redshift Survey (2QZ) for comparison. A Kolmogorov-Smirnov test does not distinguish any of the distributions. Note that we plot all QSOs in this plot, including those with a questionable redshift, and that the highest-redshift QSO in the WYFFOS part of our survey has a questionable identification. We only include WYFFOS objects observed within 20 arcminutes of the plate centre.

| Field ($\alpha$ J2000) | | OC* | IDs binned by $b_J$ as QSO/non-QSO/no-ID | | | | |
|---|---|---|---|---|---|---|---|
| $h$ | $m$ | % | 20.75 | 21.25 | 21.75 | 22.25 | Total |
| 12 | 30 | 36.0 | 4/1/3 | 0/5/6 | 5/2/9 | 0/0/9 | 10/8/27 |
| 12 | 50 | 41.7 | 0/2/5 | 2/0/6 | 7/3/14 | 1/1/9 | 10/6/34 |
| 13 | 44 | 26.9 | 0/2/2 | 1/2/1 | 4/0/3 | 2/1/10 | 7/5/16 |
| 17 | 19 | 26.4 | 3/2/5 | 3/1/6 | 4/2/9 | 0/0/3 | 11/5/23 |
| 17 | 31 | 35.8 | 2/2/4 | 3/6/8 | 2/1/10 | 1/0/3 | 8/10/25 |
| Total | | 33.2 | 9/9/19 | 9/14/27 | 22/8/45 | 4/2/34 | 46/34/125 |
| 12 | 30 | 36.0 | (3/1/4) | (0/5/6) | (3/2/11) | (0/0/9) | (7/8/30) |
| 12 | 50 | 41.7 | (0/2/5) | (1/0/7) | (5/3/16) | (0/1/10) | (6/6/38) |
| 13 | 44 | 26.9 | (0/2/2) | (1/2/1) | (2/0/5) | (1/1/11) | (4/5/19) |
| 17 | 19 | 26.4 | (2/2/6) | (2/1/7) | (3/2/10) | (0/0/3) | (8/5/26) |
| 17 | 31 | 35.8 | (2/2/4) | (3/6/8) | (1/1/11) | (1/0/3) | (7/10/26) |
| Total | | 33.2 | (7/9/21) | (7/14/29) | (14/8/53) | (2/2/36) | (32/34/139) |
| $0.3 < z < 2.2$ | | | 7(5) | 8 (6) | 14(8) | 4(2) | 35(23) |

Table 6.11: As for Table 6.10, for the WYFFOS section of the survey described in this chapter. The area of each individual field is 0.35 degrees (1.75 square degrees in total for the 5 fields). *"OC" denotes observational completeness - the fraction of the total number of QSO candidates in a field that were observed with a fibre. We only include objects observed within 20 arcminutes of a WYFFOS plate centre.

our redshift completeness, we can check that we sample a similar range of redshift to the 2QZ. In Fig. 6.13, we plot the redshift distributions of our surveys using WYFFOS and the 2dF Spectrograph in comparison to the 2QZ. A Kolmogorov-Smirnov test cannot reject the hypothesis that all 3 of the distributions in Fig. 6.13 are drawn from the same parent population of QSOs. If anything, perhaps our samples contain fewer QSOs at $z < 0.6$, which may be symptomatic of the star-galaxy separation used to tag stellar objects in the SDSS EDR (Stoughton et al. 2002) being more restrictive than that used in creating the 2QZ input catalogue (Smith et al. 1997). The single QSO in our WYFFOS sample at $z > 3$ has a questionable identification (i.e. is marked with a "?" in our lists of objects above).

Throughout the rest of this chapter, we shall consider a number of different QSO samples. Firstly, we shall assume our QSOs are a representative sample of all QSOs in the range $0.3 < z < 2.2$. Secondly, we shall often split our samples into"good" identifications (which are only those listed without a question mark in our lists of objects, above) and "all" identifications (which include those objects with questionable identifications, i.e. including those objects listed with a question mark in our lists of objects, above). We shall often refer to those objects that make the subset of "all" objects but not the subset of "good" objects as objects with a "questionable redshift" or a "questionable identification".

By "lists of objects, above" in this paragraph, we mean Tables 6.3, 6.4, 6.5, 6.6, 6.7, 6.8 and 6.9.

Throughout our analysis of the QSO number-magnitude relation, we shall only discuss fields that were observed for at least 2.5 hours, as the incompleteness of fields with a shorter integrated exposure time restricts any statistical analysis of these fields. Before discussing the QSO number-magnitude relation, we need to define the completeness of our surveys. In keeping with the patois used in this thesis (see Chapter 2), we shall refer to the fraction of QSO candidates observed with a fibre as the *observational completeness* and the fraction of objects identified from the sample of QSO candidates that were observed with a fibre as the *spectroscopic completeness*. In Tables 6.10 and 6.11 we list objects observed with the 2dF spectrograph and WYFFOS, respectively, as a function of magnitude. The subset of objects with good quality identifications are listed in parentheses. We list objects in three categories; QSOs, non-QSOs (stars and galaxies) and objects that were observed but had no identification. We also list the subset of QSOs in the range $(0.3 < z < 2.2)$. Note that for those objects observed with WYFFOS, we continue only to consider objects within 20 arcminutes of the plate centre.

In Fig. 6.14 and Fig. 6.15, we show the spectroscopic completeness as a function of magnitude for the 2dF and WYFFOS parts of our survey. The filled points in these plots are for all objects, irrespective of identification quality. The open points are the same fields but only considering good identifications. The lines mark the completeness taken across all fields (i.e. the rows labelled "Total" in Tables 6.10 and 6.11). Obviously, the completeness is much higher when we consider all identifications of any quality, particularly in the case of our 2dF data. The higher completeness of all objects in our 2dF survey is probably due to the longer exposure times. Additionally, the software, *autoz*, used to assign initial identifications to the 2dF objects is probably better than the eye at identifying possible stars from low signal-to-noise spectra. However, the larger exposure time used in our 2dF survey (as compared to our WYFFOS survey) is still not sufficient to identify faint stars with certainty, hence our 2dF survey completeness drops significantly when only the good identifications are considered. If we only consider good identifications, both the 2dF and WYFFOS parts of our survey are, spectroscopically, about 40 per cent complete in the range $21 \leq b_J < 22$. The similarity in completeness of the good identifications in our WYFFOS and 2dF surveys is due to the fact that mainly QSOs and NELGs populate the good samples, which are easier objects to identify. No objects were observed to $b_J > 22$ in the 2dF part of our survey, which, in retrospect, seems to have been a wise decision, as

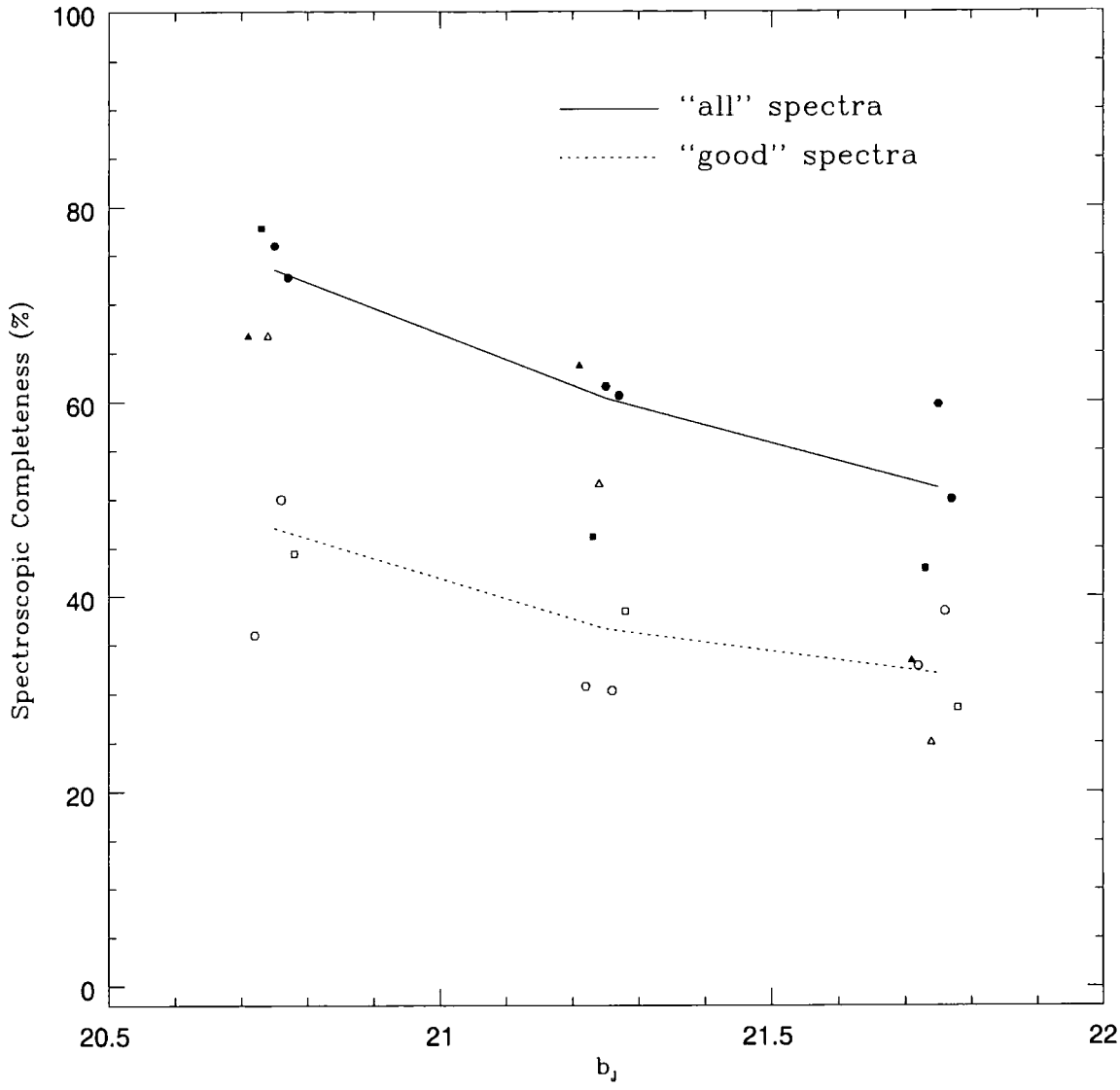


Figure 6.14: We display The spectroscopic completeness of the 2dF part of our survey for each individual plate in bins of $b_J$ magnitude. We define the spectroscopic completeness as the fraction of objects observed with a fibre that were positively identified as a star, galaxy or QSO. Filled points represent all objects, irrespective of the quality of the identification. Open points mark the stricter criterion of only considering objects with absolutely certain identifications as positively identified (i.e. not considering objects marked with a "?" in previous tables to have been positively identified). The solid line marks the combined result across all fields for all objects. The dashed line marks the combined result across all fields for objects with only absolutely certain identifications.

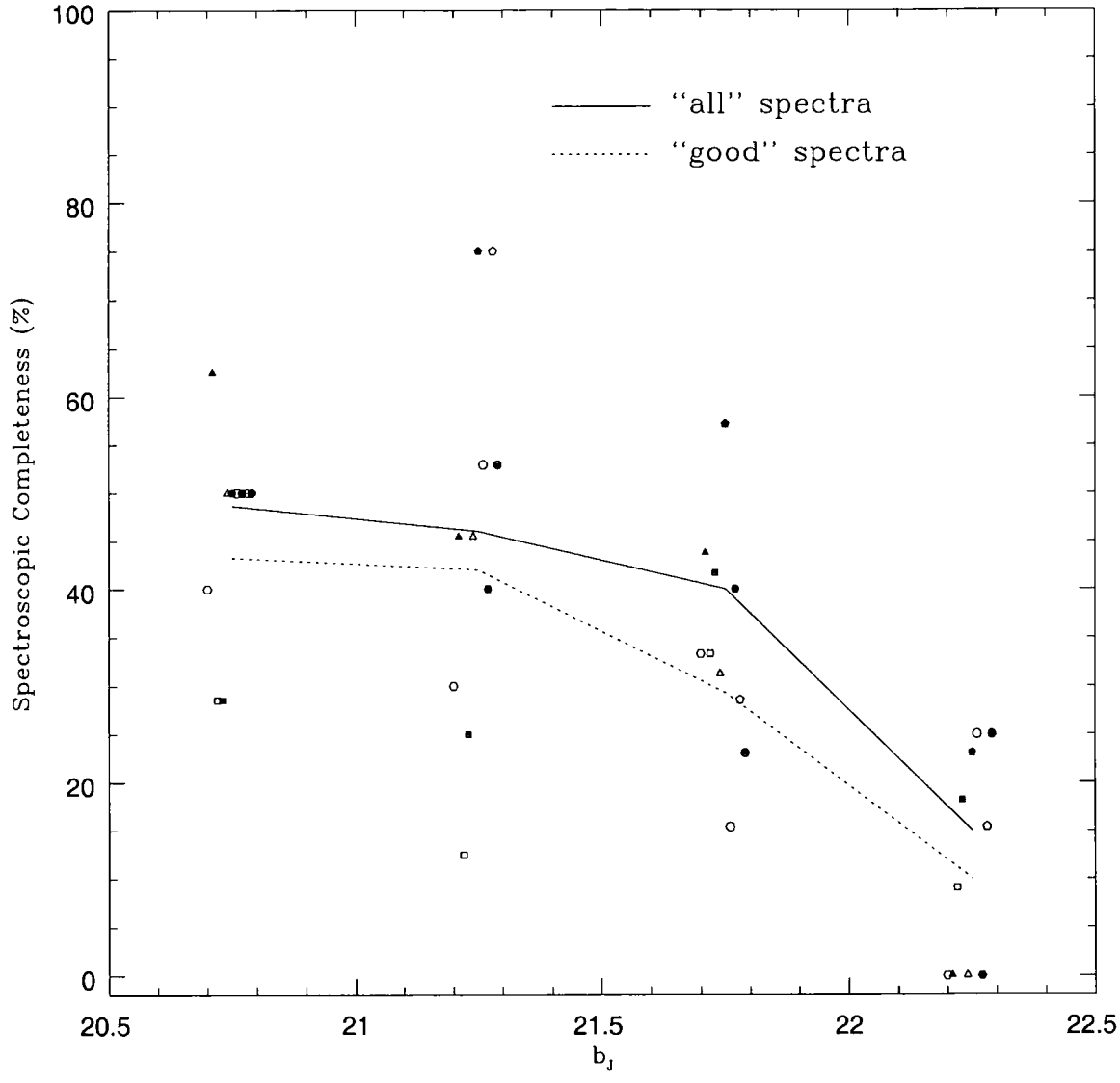


Figure 6.15: As for Fig. 6.14 but for the WYFFOS part of our survey. We only include WYFFOS objects observed within 20 arcminutes of the plate centre.

the spectroscopic completeness of our WYFFOS survey drops significantly in the faintest bin, to about 10 per cent.

In Fig. 6.16, we display the (differential) QSO number-magnitude counts (in the redshift range $0.3 < z < 2.2$) for the 2QZ, the surveys of Koo & Kron (1988) and Boyle, Jones & Shanks (1991), and our WYFFOS and 2dF surveys. We have corrected our data for dust, observational incompleteness and spectroscopic incompleteness. We have corrected for spectroscopic incompleteness on each individual field in each magnitude bin. Note that in correcting for incompleteness we assume that the identified samples are characteristic of the samples as a whole, which may not be a fair assumption if, say, more stars remain unidentified in a given spectroscopic sample than QSOs (i.e. QSOs are more likely to be identified at a given magnitude than stars). We plot four different subsets of data from our surveys. Triangles mark the counts of objects with good identifications in the WYFFOS and 2dF parts of our survey individually. Filled and open cirles compare the combined data from our WYFFOS and 2dF surveys when only good identifications are considered and when all identifications are considered, respectively. The 2dF data have about 3 times as many QSOs contributing as the WYFFOS data, and dominate the combined set of data. Note that all of our data are consistent, irrespective of whether the contributing QSOs were observed with 2dF or WYFFOS, or whether we consider the best sample of objects or all objects, although the good identifications project a marginally larger number of QSOs. This large number of QSOs projected from fewer objects is perhaps slightly counter-intuitive but arises simply because the correction for spectroscopic incompleteness is larger. We do not plot points fainter than $b_J = 22$ as there are very few QSOs in this bin and small-number statistics will dominate. Note that the faintest bin displayed in Fig. 6.16 for our 2dF survey may be plotted 0.1-0.2 magnitudes too faint, as our QSO *candidates* weren't corrected for dust before being limited to $b_J = 22$, although the increased spectroscopic incompleteness in this bin will mollify the effect. In any case, our 2dF and WYFFOS surveys remain consistent in the faintest bin.

Some of our points at $b_J < 21$ are inconsistent with the 2QZ. This may be because the 2dF has completeness problems in its faintest bin but is more likely to be due to the fact that when targets were selected for both our 2dF and WYFFOS surveys, objects in the range $21 \leq b_J < 22.2$ may have consciously been given priority over brighter objects if there was a conflict when assigning a fibre. In fact, checking the observational completeness as a function of magnitude of our two surveys, we find that a single correction for observational incompleteness is fair in the case of our WYFFOS survey, however, in

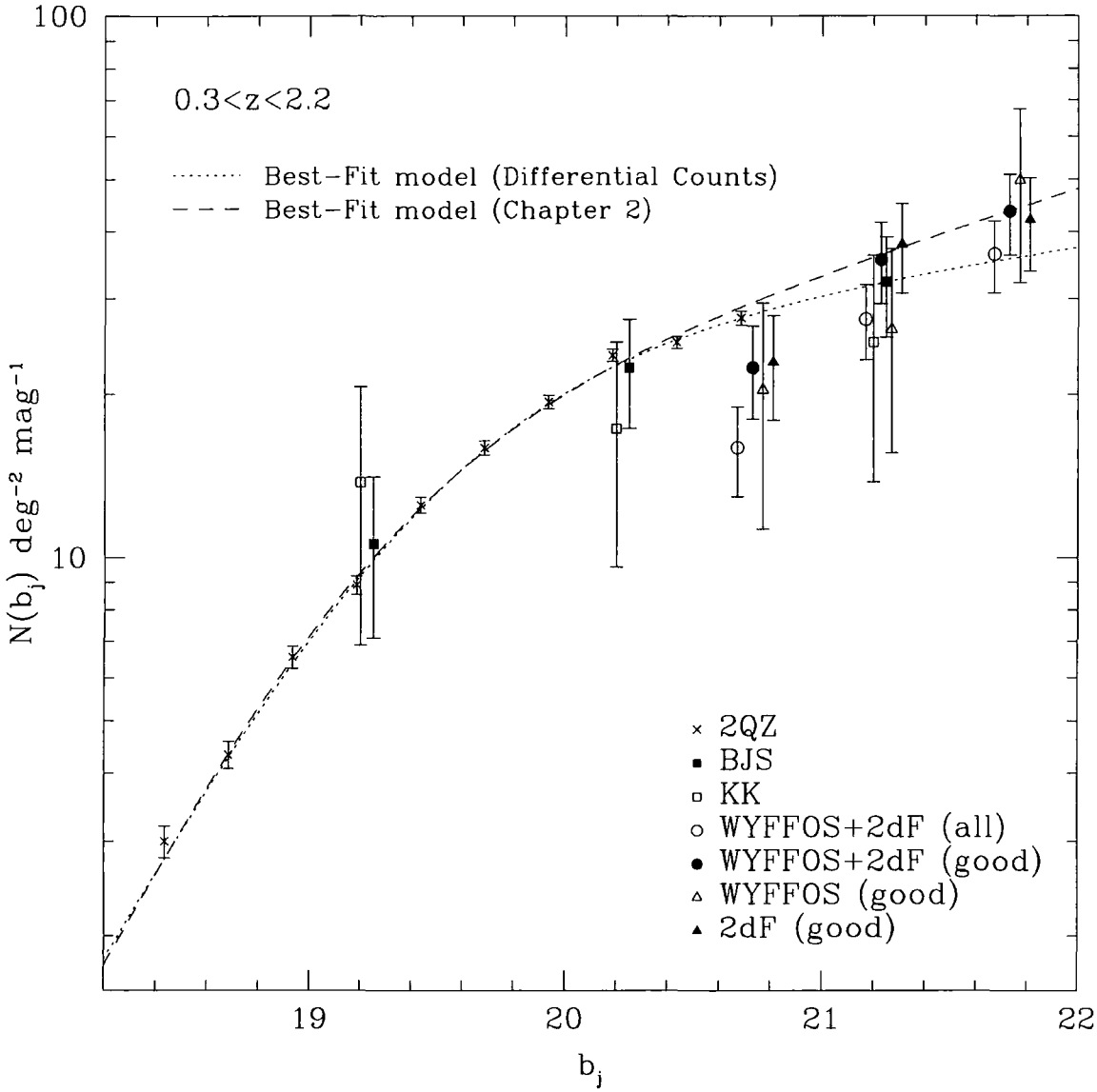


Figure 6.16: QSO numbers in bins of $b_J$ for QSOs in the range $0.3 < z < 2.2$. Plotted are points from the 2QZ (Croom et al. 2003), and from 2 other surveys, marked BJS (Boyle, Jones & Shanks 1991) and KK (Koo & Kron 1988). Several sets of points are included from our WYFFOS and 2dF surveys: Individual 2dF and WYFFOS survey numbers for definite QSOs, and combined 2dF and WYFFOS numbers for both definite QSO identifications ("good") and for all QSO identifications ("all"), irrespective of spectroscopic quality (i.e. also objects marked "?" in above tables). Our data are corrected for observational and spectroscopic incompleteness and obscuration by dust. Two models are displayed, the dashed line is a fit to the 2QZ number-counts from Section 2.3 of Chapter 2, the dotted line is a fit to the (differential) counts in this plot, including the 2QZ points (the crosses) and our "good" combined data (the filled circles). Some data are offset to prevent points merging. Errors in this plot are Poisson.

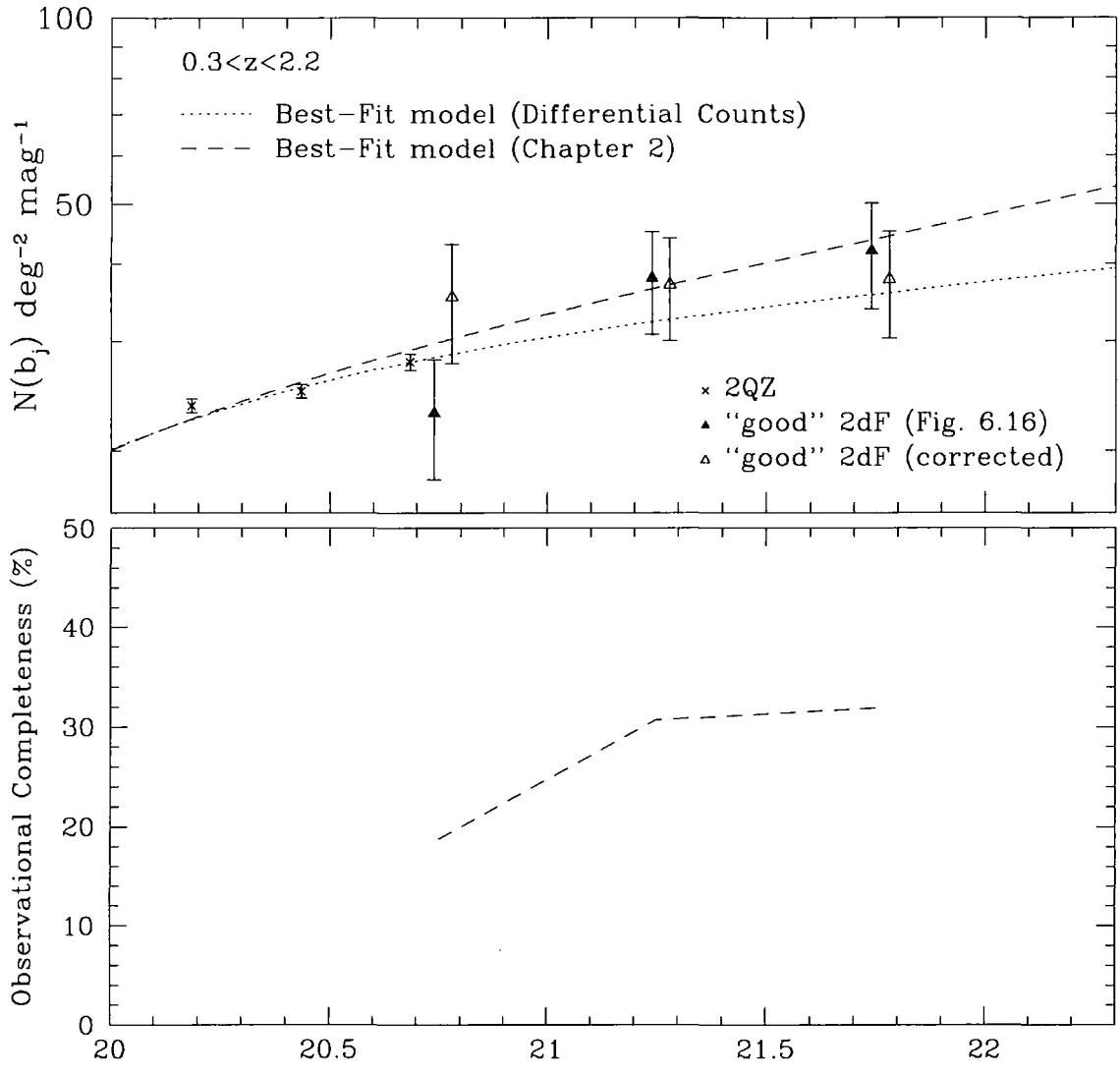


Figure 6.17: In the lower panel we illustrate the lower observational completeness at the bright end of our survey with 2dF. This is unexpected but almost certainly due to subconsciously targeting fainter QSO candidates when placing survey fibres. In the upper panel we show the effect of this lower observational completeness on the QSO number-counts displayed in Fig. 6.16. The closed triangles show the "good" 2dF data from Fig. 6.16. The open triangles show the number-counts when accounting for the magnitude dependence of the observational completeness. Errors in this plot are Poisson.

the case of our 2dF survey, the observational completeness is significantly lower in the brightest bin. We illustrate this point in Fig. 6.17, where (in the lower panel) we plot the observational completeness as a function of magnitude and in the upper panel we show what correcting for this magnitude-dependent function of observational completeness does to the number-count data. Adopting magnitude-dependent observational completeness doesn't influence our model-fits to the combined data, as the faint 2dF data are consistent whether we adopt a constant correction for observational completeness or not and statistics in the brightest bin are dominated by the 2QZ data. Note that the WYFFOS counts, as plotted in Fig. 6.16 are genuinely low in the brightest bin, which may simply be due to small-number statistics.

We also plot two models in Fig 6.16. The dashed line is the SPL model from Section 2.3 of Chapter 2, representing the best compromise fit to both the integral and differential 2QZ QSO number-counts. The dotted line is the minimal $\chi^2$ fit to both the 2QZ differential counts and the data points combined from the subset of our 2dF and WYFFOS surveys with good identifications (the filled circles in Fig. 6.16). The dashed line is an excellent fit to our faint data points, as well as being the best compromise fit to the 2QZ differential and integrated counts. The best-fit model to the 2QZ differential counts underestimates our counts fainter than $b_J = 21$ (but is still the best-fit model, as our points have practically no statistical power compared to the 2QZ). We can integrate up the models displayed in Fig. 6.16 and determine the faint-end slope of the integrated counts over the range a magnitude fainter than the limit of the 2QZ, as in Chapter 2. The fit from Chapter 2 has a faint-end slope of 0.29. The fit to the differential counts, which is inconsistent with our faint data at around the $1\sigma$ level, has a faint-end slope of 0.26, suggesting a value for the slope of the QSO number-counts a magnitude fainter than the limit of the 2QZ of $0.29 \pm 0.03$, which, as it turns out, was the value adopted throughout this thesis (although we have improved the error-bar).

### 6.4.3 The 2dFSDSS QSO Survey

The intended purpose of our 2dF and WYFFOS surveys was to better determine the slope of the integrated QSO number-magnitude counts fainter than the limit of the 2QZ. We confirm a value of $0.29 \pm 0.03$ for this slope, in close agreement with the value adopted in Chapter 2. The quoted error bars, however, represent a stronger limit than before, as our combined 2dF and WYFFOS data sets contain around 3 times as many QSOs in the range $21 \leq b_J < 22$ as the surveys of Boyle, Jones & Shanks (1991) and Koo & Kron

(1988) combined, which were used in Chapter 2 to check the form of the faint-end QSO integrated number counts. However, our surveys still only contain about 70 absolutely certain QSOs in the range $21 \leq b_J < 22$ and are only about 40 per cent spectroscopically complete in this magnitude range. The lack of spectroscopic completeness is a particular problem and might cause us to overestimate the number of QSOs at a given magnitude, as in correcting up for spectroscopic incompleteness we assume that the identified objects are characteristic of the sample as a whole, which would not be the case if, say, stars are less likely to be identified than QSOs. Thus, there is still scope for a larger, more refined QSO survey to magnitudes fainter than the 2QZ. Our surveys suggest that this is certainly possible to $b_J < 22$ using the 2dF spectrograph on the AAT. The easiest way to improve the completeness of any such QSO survey would be to try to target more QSOs and fewer stars when selecting candidates. Our faint sample of QSOs, which is larger than any sample in the literature at magnitudes fainter than $b_J = 21$, is thus particularly useful in seeking a new UVX algorithm to target QSOs while rejecting stars. In fact, members of the SDSS and 2dF QSO survey teams are collaborating to construct just such a faint QSO survey. The Principal Investigators in this collaboration are Prof. T. S. Shanks (UK) and Dr. S. M. C. Croom (Australia). An algorithm to target QSOs by colour has been constructed by G. T. Richards (of the University of Pennsylvania) with help from P. J. Outram (of the University of Durham). In this Section, we will use our faint QSO data to test the efficacy of this algorithm. We shall refer to this algorithm, which uses up-to-date SDSS ($u, g, r, i$ and $z$) magnitudes to target QSOs for follow-up with 2dF, as *the 2dFSDSS algorithm*. We will consider QSOs identified in both our WYFFOS and 2dF surveys.

Objects identified in our 2dF and WYFFOS surveys were cross-referenced against the SDSS Data Release 1 (Abazajian et al. 2003), which contains the most up-to-date SDSS imaging. The best available (PSF) SDSS magnitudes for each of the objects identified in this chapter were derived by matching the position of objects in our 2dF and WYFFOS surveys (which all had positions taken from the SDSS EDR) to the nearest object in the SDSS Data Release 1 (henceforth DR1). These magnitudes were then corrected for absorption in the SDSS passbands using the dust maps of Schlegel, Finkbeiner & Davis (1998), as the 2dFSDSS algorithm makes colour cuts on the basis of dust-corrected magnitudes. Over 95 per cent of our objects have a match in the SDSS DR1 within 2 arcseconds. Across all fields in our 2dF and WYFFOS surveys, irrespective of exposure time, there is either a good or a questionable identification for 174 QSOs and 140 non-

QSOs. Initially, the 2dFSDSS algorithm rejects objects with large magnitude errors; greater than 0.4 in $u$, 0.13 in $g$ or $r$, 0.2 in $i$ or 0.6 in $z$. Of our identified objects, 13 QSOs and 15 non-QSOs would be rejected by the 2dFSDSS algorithm due to large errors in one of their PSF magnitudes.

The next step in the 2dFSDSS algorithm is to reject objects that have magnitudes that place them in regions of $u, g, r, i, z$ parameter space dominated by Main Sequence stars (see, e.g., Richards et al. 2002). We illustrate this part of the algorithm in Fig. 6.18. It turns out that none of our objects are rejected by this part of the 2dFSDSS algorithm, which might be expected, as our UVX selection of QSO candidates outlined in Section 6.2 was also designed to reject the main stellar locus. Ultimately then, 161 QSOs and 125 non-QSOs are accepted by the algorithm up to this point, and so far, only those objects with poor magnitude determinations in at least one band in the SDSS DR1 have been rejected.

Once the Main Sequence stellar locus is rejected, the 2dFSDSS algorithm makes a few final refinements to the parameter space to specifically select areas that only QSOs populate. The first refinement is essentially the typical UVX selection method discussed in Section 6.2 but in $(g-r,\, r-i)$ space. The second refinement is the result of searching for a simple refinement in SDSS DR1 colour-colour space that rejects faint stars but not faint QSOs, while maintaining a high level of completeness out to a redshift of $z \sim 2$. This cut, in the $(u-g,\, g-i)$ plane is the key difference between the UVX selection method used in Section 6.2 to find QSO candidates and the candidate selection for the 2dFSDSS Survey. The exact form of the final colour cuts is slightly dependent on the $g$ magnitude of the candidate QSOs. Slightly stricter cuts are used for QSO candidates with $g \geq 21.15$, to try and increase the completeness of the faint sample. The bright limit of the 2dFSDSS survey is currently intended to be $g \geq 18.00$.

In Fig. 6.19 (for $g < 21.15$) and Fig. 6.20 (for $g \geq 21.15$) we plot colours of the 161 QSOs identified in our WYFFOS and 2dF surveys (and thus far accepted by the 2dFSDSS algorithm) in the $(g-r,\, r-i)$ and $(u-g,\, g-i)$ planes. We plot good identifications as filled shapes and questionable identifications as open shapes. The colour cuts imposed by the 2dFSDSS selection algorithm are plotted as dashed lines. The 2dFSDSS algorithm would accept an object as a QSO candidate if it lies in the parameter space to the left of these dashed lines (in both panels of a given plot). Of the 161 QSOs, 82 have a dust-corrected magnitude in the range $18.00 \leq g < 21.15$ and 79 have $g \geq 21.15$. If we only consider good identifications in our WYFFOS and 2dF samples, we consider 68 QSOs

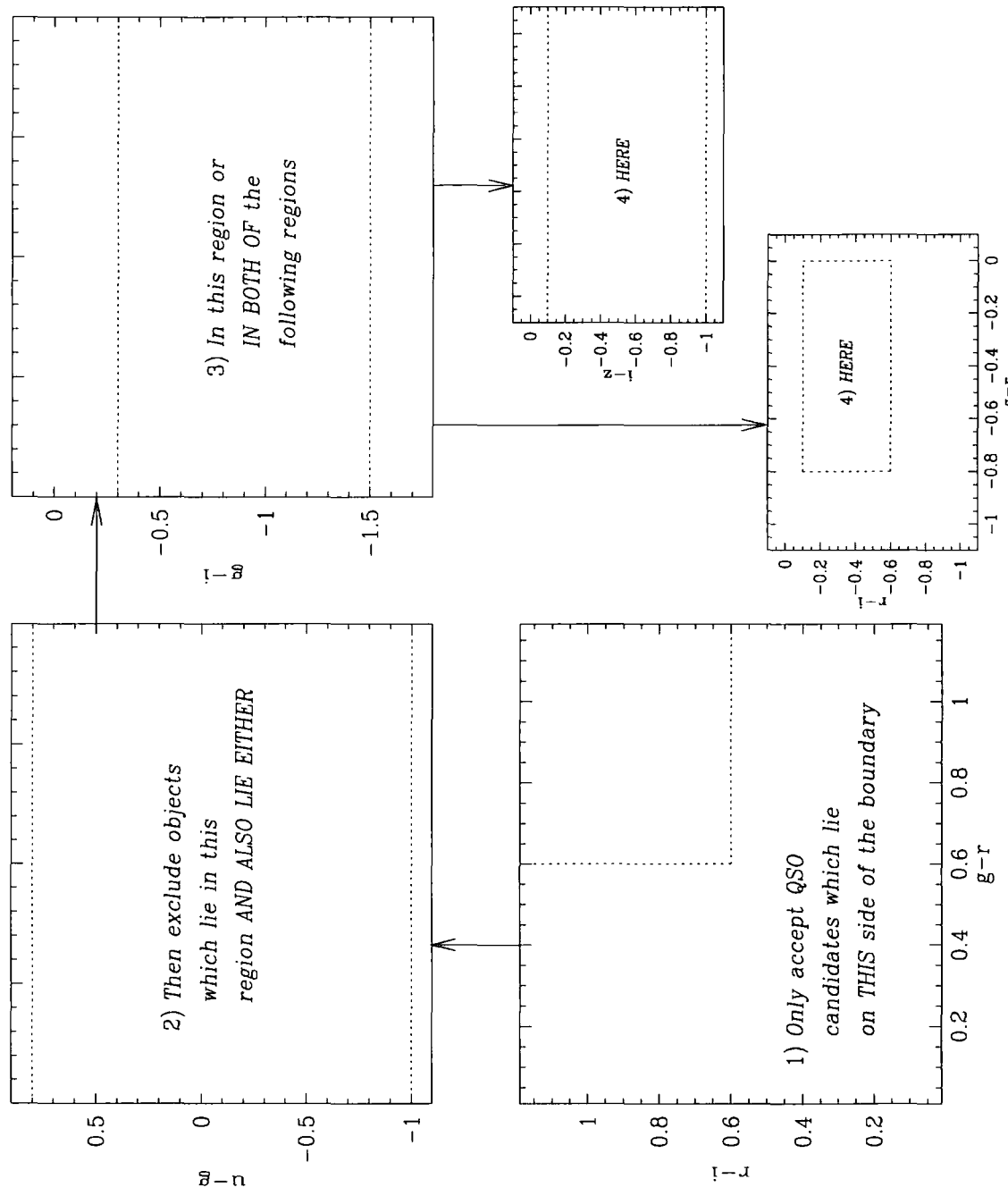


Figure 6.18: This figure illustrates the inital colour cuts used to reject Main Sequence and Blue Horizontal Branch stars when picking faint QSO candidates out of the SDSS to observe as part of the 2dFSDSS QSO Survey. All colours illustrated in the plot are from the SDSS magnitude system (Fukugita et al. 1996). *The algorithm used for finding QSO candidates to be included in the 2dFSDSS QSO Survey is courtesy of P.J. Outram (Durham University) and G.T. Richards (Pennsylvania State University).*

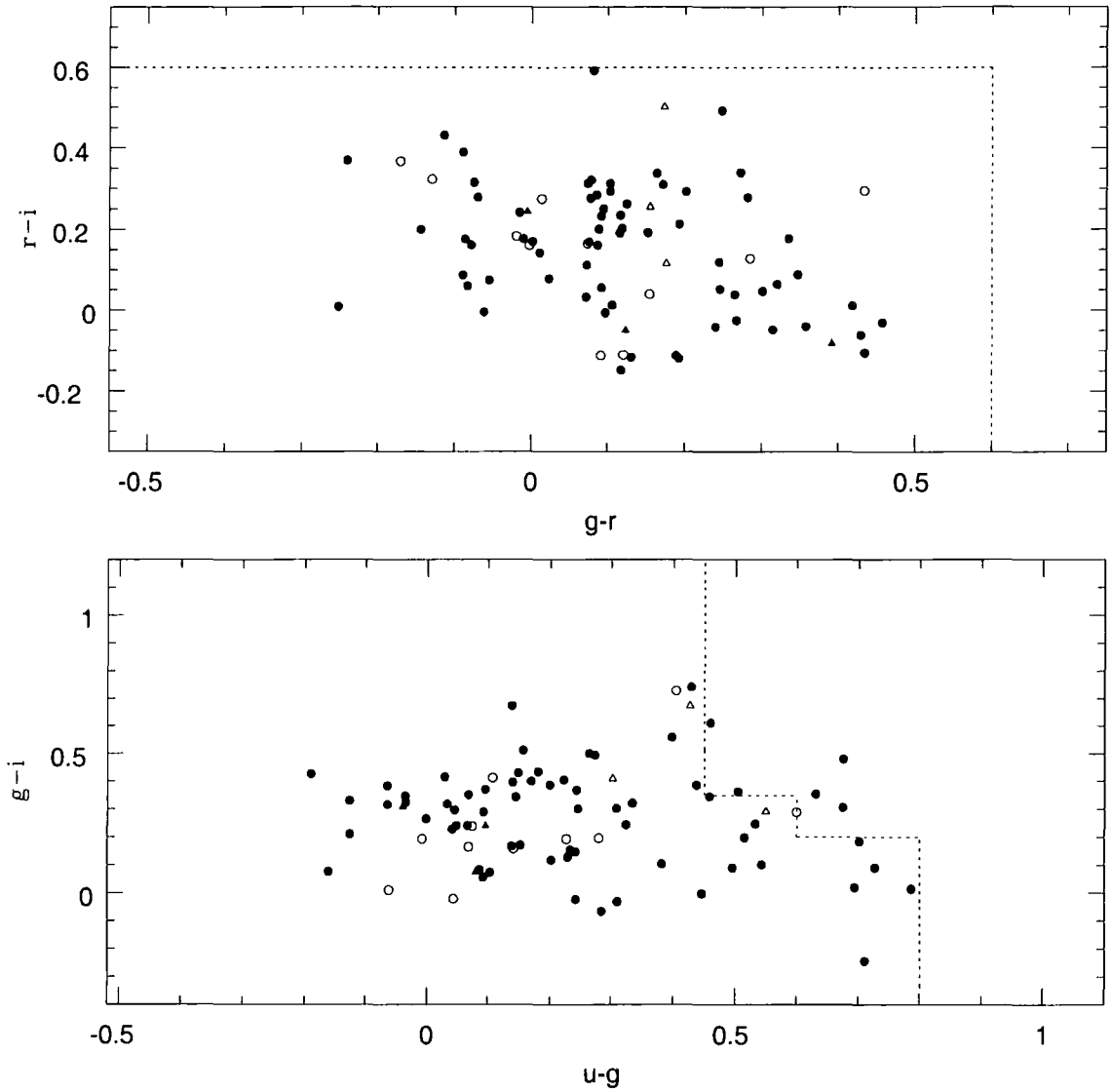


Figure 6.19: The 2dFSDSS QSO Survey selects QSO candidates using a colour selection algorithm developed by P.J. Outram (Durham University) and G.T. Richards (Pennsylvania State University). After a series of initial cuts are made on the basis of SDSS colours, magnitude errors and star-galaxy separation (see Fig. 6.18), objects with magnitudes in the range $18.00 \leq g < 21.15$ are tagged as potential QSOs if they lie leftwards of cuts in the $(g-r,\ r-i)$ and $(g-i,\ u-g)$ planes illustrated by dotted lines in this diagram. We plot QSOs identified in the WYFFOS and 2dF surveys outlined in this chapter, which meet the initial selection criterion and lie in the range $18.00 < g < 21.15$. UVX QSOs with a good identification are plotted as filled circles. UVX QSOs with questionable identifications (i.e. those marked "?" in above tables) are plotted as open circles. Also plotted, as triangles, are colours of the X-ray loud QSOs observed with WYFFOS. In all, about 95 per cent of the QSOs represented would be accepted as QSO candidates for the 2dFSDSS Survey.

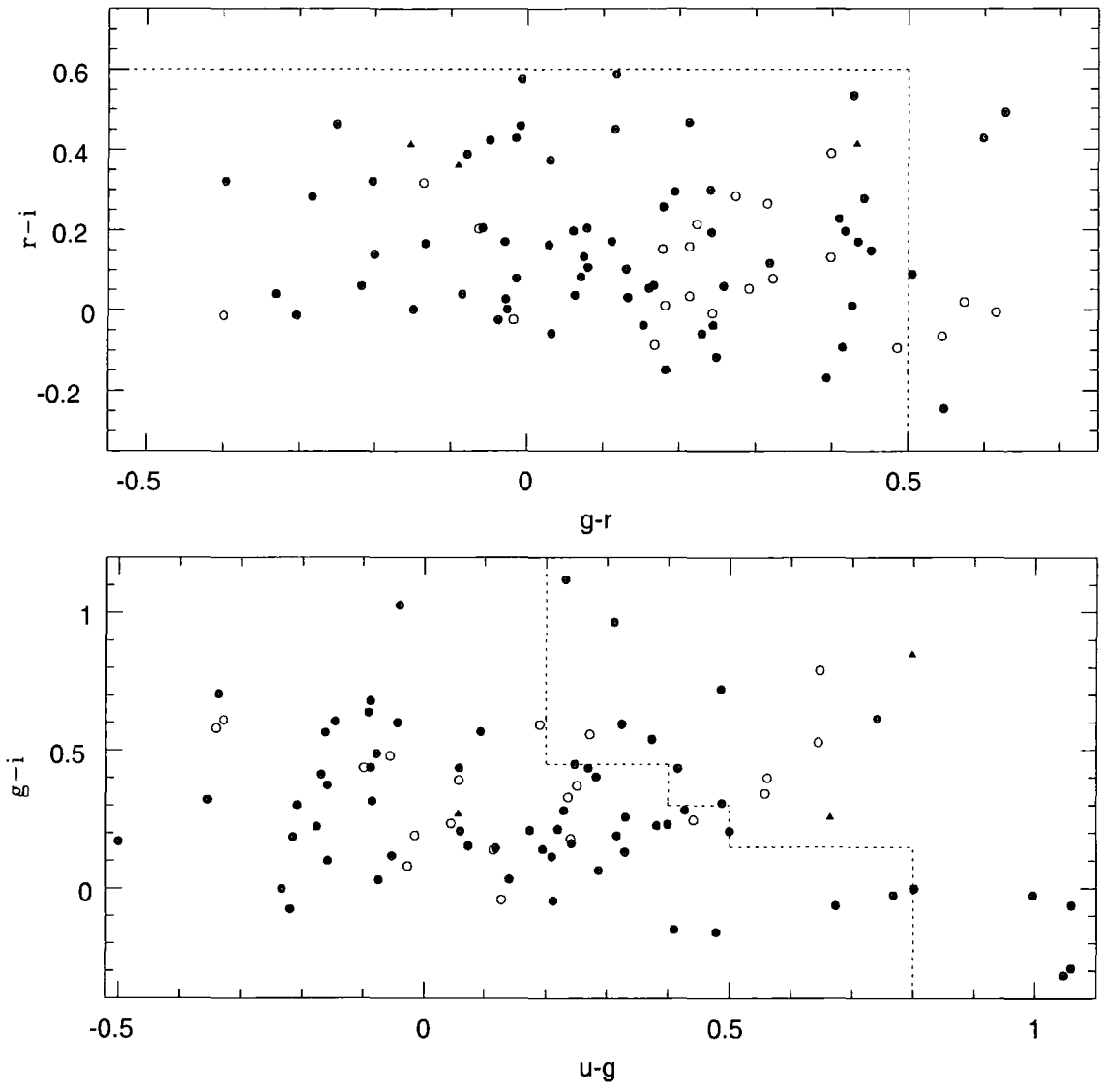


Figure 6.20: As for Fig. 6.19 but for QSOs in the magnitude range $21.15 \leq g < 21.85$. Note that the dotted lines (which mark the boundary leftwards of which objects would be tagged as QSO candidates) trace a slightly different parameter space than in Fig. 6.19. In all, about 70 per cent of the QSOs represented in this plot would be accepted as QSO candidates for the 2dFSDSS Survey.

with $g < 21.15$ and 59 QSOs with $g \geq 21.15$. Very few of our QSOs would be rejected by cuts in the $(g-r, r-i)$ plane, which we would expect, as our QSOs are almost all already UVX selected (we also consider the X-ray loud QSOs targeted with WYFFOS). Slightly more QSOs would be rejected as 2dFSDSS candidates by cuts in the $(u-g, g-i)$ plane. In all 78/82 of our $g < 21.15$ QSOs would be accepted by the 2dFSDSS algorithm (64/68 for the subset with good identifications) and 52/79 of our $g \geq 21.15$ QSOs would be accepted (41/59 for good identifications). If our QSO population is generally representative of faint QSOs, then we would expect the 2dFSDSS survey to include 95 per cent of QSOs with dust-corrected magnitude $g < 21.15$ and 65-70 per cent of QSOs with dust-corrected magnitude $g \geq 21.15$.

So, the 2dFSDSS algorithm does a reasonable job of targeting UVX QSOs fainter than $g = 21.15$ and an excellent job of targeting UVX QSOs brighter than $g = 21.15$. What level of contamination by non-QSOs might be expected? In Fig. 6.21 (for $g < 21.15$) and Fig. 6.22 (for $g \geq 21.15$) we plot the position of the 125 non-QSOs identified in our WYFFOS and 2dF surveys (that have accurate PSF magnitudes in the SDSS DR1). Again we note that objects would be accepted as QSO candidates if they lie in the parameter space to the left of the dashed lines in both panels of the figure they are plotted in. Of the $g < 21.15$ non-QSOs, 11/68 would be targeted as QSO candidates in the 2dFSDSS Survey (6/27 for good identifications). Of the $g \geq 21.15$ non-QSOs, 13/57 would be targeted in the 2dFSDSS Survey (9/22 for good identifications). We thus project that the 2dFSDSS Survey will contain only 16-22 per cent of $g < 21.15$ non-QSOs that would be contained in a UVX sample and 23-41 per cent of the $g > 21.15$ non-QSOs that would be contained in a UVX sample. Overall, we project (having scaled the relative numbers of QSOs and non-QSOs) that 85 per cent of $g < 21.15$ QSO candidates will turn out to be QSOs (89 per cent considering only good identifications) and 76 per cent of $g \geq 21.15$ QSO candidates will turn out to be QSOs (78 per cent if we only consider good identifications). These are favourable projected QSO fractions when we consider that 58 per cent of objects positively identified in the 2QZ are QSOs and 50 per cent of objects positively identified in our WYFFOS and 2dF surveys are QSOs. It is very difficult to find a single colour-cut better than the $(u-g, g-i)$ cut the 2dFSDSS survey uses to target QSOs - at least using our 2dF and WYFFOS data.

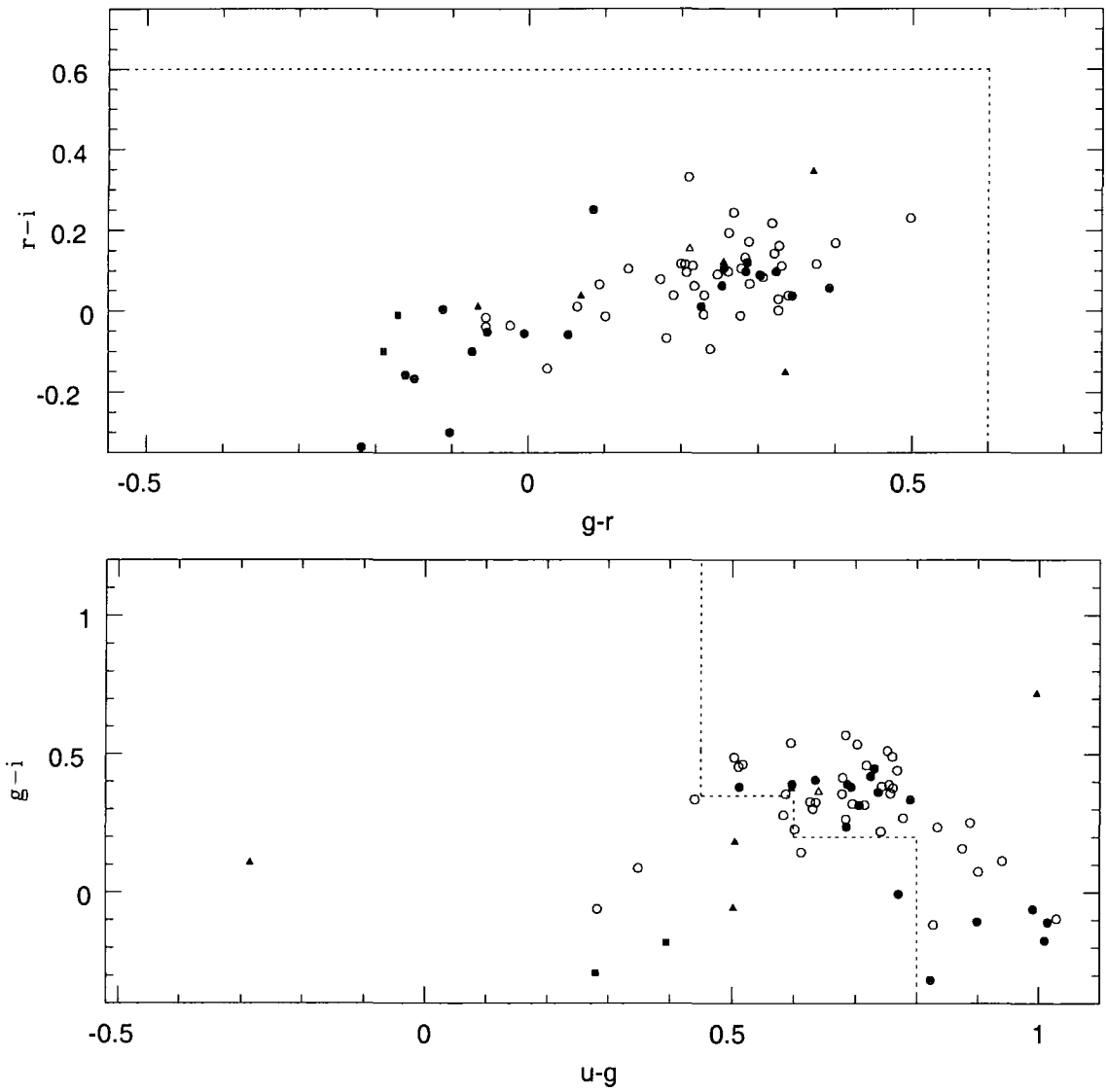


Figure 6.21: As for Fig. 6.19 but for non-QSOs in the magnitude range $18.00 \leq g < 21.15$. Stars are represented by circles (open shapes represent objects with questionable identifications). NELGs (and LINERs) are marked by triangles. White Dwarfs are represented by squares. Objects to the left of the dashed lines would be tagged as candidate QSOs. In all, about 80 per cent of the objects represented in this plot would be rejected as QSO candidates for the 2dFSDSS Survey.

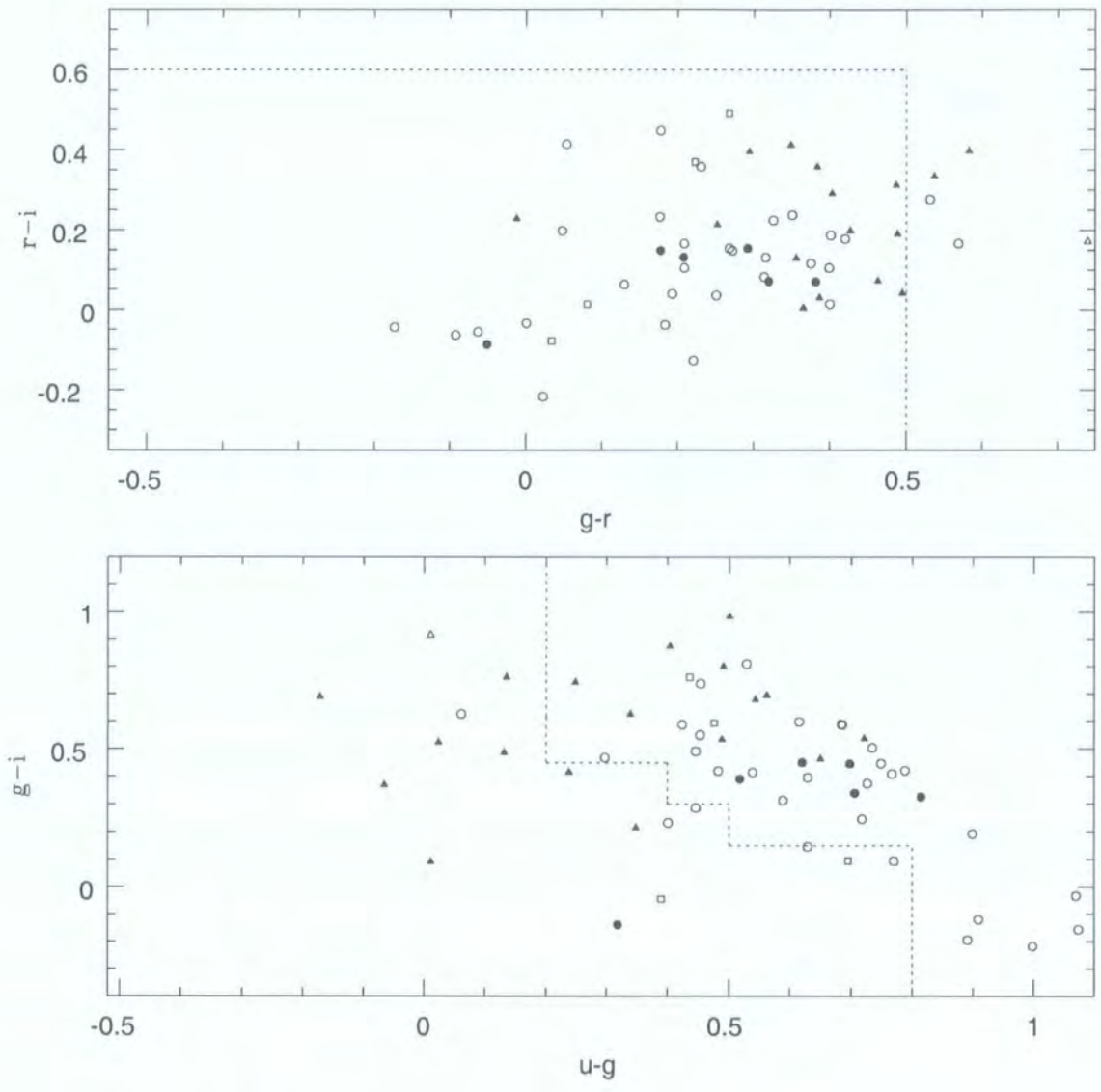


Figure 6.22: As for Fig. 6.20 but for non-QSOs in the magnitude range $21.15 \leq g < 21.85$. Stars are represented by circles (open circles being the questionable identifications). NELGs (and LINERs) are marked by triangles. White Dwarfs are represented by squares. Objects to the left of the dashed lines would be tagged as candidate QSOs. In all, about 70 per cent of the objects represented in this plot would be rejected as QSO candidates for the 2dFSDSS Survey.

## 6.5 Conclusions

We have conducted two surveys targeting QSOs in the range $20.5 \leq b_J < 22.5$. We have used SDSS EDR photometry to select a UVX sample of QSO candidates similar to that targeted in the 2dF QSO Redshift Survey and shown that this photometry is likely accurate to $\pm 0.1$ mag in $b_J$. These QSO candidates were observed with WYFFOS on the WHT and the 2dF Spectrograph on the AAT over a total of 7 nights in March and May 2002. In total, 5.5 nights of useful observing time resulted, allowing useful total exposures of $\sim$ 3 hours for 5 WYFFOS fields and $\sim$ 6 hours for 4 2dF fields. Some additional WYFFOS fields were exposed for 1-2 hours. One WYFFOS field was observed in conjunction with a second project targeting X-ray loud QSO candidates. All of the 2dF fields were observed in conjunction with a second project. The candidate QSO spectra were reduced using various IRAF packages, in the case of the WYFFOS data, and the dedicated 2dF reduction package *2dfdr* in the case of the 2dF data. In total, 163 UVX QSOs were identified fainter than $b_J = 20.5$. Of these, 103 were observed with 2dF to $b_J = 22$ and 60 with WYFFOS to $b_J = 22.5$. Of the QSOs observed with WYFFOS, 2 have low-quality spectra due to flat-fielding problems (although are almost certainly QSOs) and a further 24 are subjectively characterised as having poor quality spectra or ambiguous redshift determinations. Of the QSOs observed with 2dF, 3 have been previously identified (Hall et al. 1996) and a further 8 are characterised as having poor quality spectra or ambiguous redshift determinations. The exact classification of what constitutes a poor spectra may not be consistent between the WYFFOS and 2dF samples, as the 2dF sample initially had identifications and redshifts assigned using an objective automated procedure, where the WYFFOS sample was entirely identified by eye, although identifications of the 2dF spectra were ultimately additionally checked by eye.

The completeness of our surveys was discussed. Notably, we have demonstrated that far fewer QSOs than expected are identified in our data sample outside of the central 40 arcminutes of a WYFFOS plate. It is well known that WYFFOS suffers some vignetting outside of the central 40 arcminute field of view (Corradi et al. 2001) and, arguing that this vignetting would severely effect our survey completeness, we decided to only consider survey statistics within the central 40 arcminutes of a WYFFOS field. The 2dF data suggested no variation in completeness with fibre position. The spectroscopic completeness was determined for both the WYFFOS and 2dF data and found to be about 40 per cent in the range $21 \leq b_J < 22$ for the strictest spectral identifications in both sets of data.

The typical observational completeness of the two surveys is about 30 per cent.

We have used our 2dF and WYFFOS data to confirm the form of the QSO number-magnitude counts in the magnitude range $21 \leq b_J < 22$ for QSOs in the redshift range $0.3 < z < 2.2$, the range for which UVX QSO samples are designed to be most complete (Véron 1983). In total, 83 of our QSOs lie in these redshift and magnitude ranges, of which 68 adhere to the strictest spectral criteria. This means our survey is 3-4 times larger than the surveys of Boyle, Jones & Shanks (1991) and Koo & Kron (1988) combined, the largest individual surveys in the literature fainter than $b_J = 20.85$. By fitting models to the differential QSO counts, we confirm the slope of the integrated QSO number-magnitude counts to be $0.29 \pm 0.03$ - essentially the value of the slope used throughout this thesis, with improved accuracy.

Finally, we have discussed the proposal of the 2dFSDSS collaboration of extending our idea of combining SDSS photometry and 2dF spectroscopy to obtain a much larger sample of faint QSOs. We have discussed the 2dFSDSS algorithm developed by G. T. Richards and P. J. Outram to improve on the UVX method of targeting faint QSOs, particularly, using a simple cut in $(u-g, g-i)$ parameter space to obtain lower rates of contamination by non-QSOs. We have estimated (using 161 QSOs and 125 non-QSOs obtained in our UVX surveys to $b_J > 20.5$) that 85-90 per cent of $g < 21.15$ candidates targeted by the 2dFSDSS algorithm will be QSOs and 75-80 per cent of $g \geq 21.15$ candidates targeted by the 2dFSDSS algorithm will be QSOs. This is a high QSO return compared to traditional UVX candidates, 50-60 per cent of which are QSOs.

# Chapter 7 | *Conclusions*

## 7.1 Summary

The mandate of this thesis was to study the statistical lensing of QSOs by foreground structure in light of the large amount of data recently made available as the 2dF QSO Redshift Survey (2QZ). In the Introduction of this thesis, we discussed the evolution of cosmology, loosely divided into the "classical era" and "modern cosmology". We noted that while the questions humans seek to ask about our Universe are eternal, the development of new theories and technologies drive the field of observational cosmology towards a more intricate understanding of the Cosmos. It could be argued, for instance, that, classically, the development of General Relativity and larger telescopes equipped for spectroscopy marked the genesis of modern cosmology. The last decade or so, has seen the wide acceptance of gravitational lensing as a tool to probe the mass (rather than the light) associated with structures in our Universe. Similarly, the last decade has been graced by the evolution of spectroscopy to the point where spectroscopic surveys of significant portions of the sky can be carried out on timescales of a few years. Still, few cosmological results have emerged from the marriage of gravitational lensing theory and these large spectroscopic surveys, and a barrage of studies addressing cosmological lensing in large surveys is likely imminent.

After introducing the classical and modern cosmological parameters that might be probed by the statistical lensing of QSOs, introducing the values of these parameters that are widely accepted as characterising the geometry of our Universe (the "Concordance Cosmology") and discussing the current standard model ($\Lambda$CDM), we turned our attention to the 2QZ. In Chapter 2, we outlined the form and content of the 2QZ and derived a number of results that would prove useful in studying properties of the QSO population as a whole. Chapter 3 reintroduced the result of Boyle, Fong & Shanks (1988), that QSOs and galaxy groups are anti-correlated and discussed the idea (Rodrigues-Williams & Hogan 1994, Croom & Shanks 1999) that this signal could be attributed to lensing by groups of galaxies diluting background QSO numbers. The groups used to study QSO-group cross-correlations were defined where galaxy numbers in the APM Survey

and the Sloan Digital Sky Survey (SDSS) are particularly dense on the plane of the sky. The lensing models used to characterise the effect of the groups' mass on QSO numbers considered the groups to be well-characterised as dark matter haloes. In Chapter 4, we investigated in what sense these groups selected on the plane of the sky might be thought to be representative of dark matter haloes in our Universe. We went on to investigate the cross-correlation between QSOs and groups selected from the 2dF Galaxy Redshift Survey (Eke et al. 2003) to best trace underlying dark matter haloes in a ΛCDM universe.

Many of the problems that arise when trying to characterise what constitues a "group" of galaxies can be circumvented by analysing the ensemble galaxy distribution, rather than groups selected from it. In Chapter 5 we considered the cross-correlation between QSOs and galaxies. Recently Gaztañaga (2003) has found a significant correlation between galaxies and QSOs taken from the SDSS, a result that might have significant implications for how mass in our Universe clusters on small scales, as compared to the light from galaxies. In Chapter 5, we checked whether the result of Gaztañaga (2003) was consistent with the distribution of QSOs in the 2QZ relative to foreground galaxies. The results of Chapter 5 (and Chapter 3) are somewhat at odds with the Concordance Cosmology, and, further, could be explained if the form of the QSO number-magnitude counts flattens significantly at magnitudes fainter than the limit of the 2QZ. In Chapter 6, we introduced two QSO surveys intended to study the form of the QSO counts to $21 \leq b_J < 22.5$. The idea of extending the 2QZ to fainter magnitudes has now been taken up by a 2dF-SDSS collaboration and our faint QSO data were useful in testing the algorithms used by this collaboration to target candidate QSOs for spectroscopy. Thus, our thesis came full circle, beginning with analysis of the completed 2dF QSO Redshift Survey and ending with preliminary analysis of what might well become the 2dFSDSS QSO Redshift Survey.

## 7.2 Main Results of This Thesis

In Chapter 2 of this thesis, we discussed the construction of angular completeness masks for the 2dF QSO Redshift Survey (2QZ), that were based on the fraction of QSOs identified compared to the number of candidate QSOs targetted by the survey. The masks make the assumption that the sample of objects not spectroscopically identified by the 2QZ contains the same fraction of QSOs as the parent catalogue. By analysing the subset of 2QZ objects that were initially classed as unidentified but were subsequently reobserved and assigned a spectral quality, we showed that the fraction of QSOs identified in the 2QZ

is likely to be representative of the fraction of QSOs in the sample of 2QZ objects that were observed without being identified. This is true even at $b_J > 20.5$, the faint end of the 2QZ. This means that correcting observed QSO numbers on the basis of the fraction of QSOs identified in the 2QZ will fairly estimate the true number of QSOs. The other main result of Chapter 2 was our analysis of the form of the integrated QSO number-magnitude counts, $N(< b_J)$. When the number of QSOs as a function of magnitude are corrected for absorption by dust in our Galaxy and for incompleteness, both the cumulative and differential data are fit very well by a Smooth Power Law (SPL) model. We found that the faint-end slope of the integrated SPL model over the range $20.85 \leq b_J < 21.85$ is $0.29 \pm 0.05$, a value consistent with the determinations of earlier authors (Boyle, Fong & Shanks 1988, Koo & Kron 1988, Boyle, Jones & Shanks 1991, Hartwick & Schade 1990).

In Chapter 3, we sought the effect of lensing by foreground groups of galaxies on QSO numbers. By cross-correlating ($z > 0.4$) 2QZ QSOs against galaxies in objectively-determined groups, we demonstrated that there is a $3\sigma$ anti-correlation between QSOs and galaxy groups of strength $\omega_{gq}(< 10') = -0.049$. The strength of this effect is consistent with the result of Boyle, Fong & Shanks (1988), with the caveat that Boyle, Fong & Shanks (1988) cross-correlated UVX QSO candidates, rather than spectroscopically identified QSOs, against galaxy groups. Care was taken to show that the anti-correlation between QSOs and galaxies in groups is not a selection effect due to the 2QZ observing process. By analysing the colours of QSOs as a function of their distance from the centres of galaxy groups, we limited the amount that dust in galaxy groups could redden QSOs, finding a 95 per cent upper limit on reddening in $b_J - r$ of 0.012, which translates to a $2\sigma$ limit on absorption in groups of $A_B < 0.04$ mag assuming the Galactic dust absorption law from Schlegel, Finkbeiner & Davis (1998), in good agreement with the limit on absorption in galaxy groups from Nollenberg, Williams & Maddox (2003). To explain the anti-correlation of QSOs and galaxy groups as an effect of absorption by dust in the groups would require an absorption of at least $A_B \approx 0.2$ mag (Boyle, Fong & Shanks 1988). We later demonstrated (see Section 5.4.4.1) that assuming even the greyest dust law measured in local galaxy groups (Calzetti, Kinney & Storchi-Bergmann 1994) would not increase our estimate of absorption in groups by more than 25 per cent.

Magnification bias (see Appendix B) predicts that where the slope, $\alpha$, of the integrated QSO number-counts, $N[< m] \propto 10^{\alpha m}$, is flatter than 0.4, an anti-correlation should be observed between QSOs and foreground structure. The strength of the expected anti-correlation depends on the exact value of the slope and the size and distribution of the

mass of foreground groups. When Croom & Shanks (1999) modelled the anti-correlation between galaxies in groups and UVX QSO candidates as an effect of statistical lensing, they weighted each group by the number of galaxies in the group, effectively skewing their result towards groups with more galaxies (and therefore higher mass groups). To study the anti-correlation between galaxy groups and QSOs in terms of statistical lensing, we remeasured the result as a cross-correlation between QSOs and group centres, with each group weighted equally. The revised anti-correlation measured was $\omega_{cq}(< 10') = -0.034$, at a significance of $3\sigma$. We then demonstrated that the anti-correlation is fit well by supposing its cause is the statistical lensing of QSOs by dark matter haloes, either NFW haloes or Singular Isothermal Spheres. Such a lensing explanation requires more mass than models with $\Omega_m = 0.3$ would suggest, at the $\sim$ 1-2$\sigma$ level, meaning the effect is somewhat at odds with the Concordance Cosmology. Alternative explanations for the anti-correlation include a combination of lensing and dust, a significant flattening of the QSO integrated number-counts ($\alpha < 0.2$ in the range $20.85 \leq b_J < 21.85$), or some combination of these factors.

A major assumption was made in the statistical lensing analysis of Chapter 3, namely, that groups selected as over-densities on the plane of the sky were representative of the dark matter haloes used to model them as lenses. In Chapter 4, we investigated how many groups defined on the plane of the sky ("2D groups") would still be considered groups if we also knew the redshifts of the galaxies in the groups ("3D groups"). As no large galaxy redshift surveys to $b_J < 20.5$ currently exist, we used a mock catalogue (Cole et al. 1998) created from an N-body simulation of dark matter in a ΛCDM universe to test our group-detection algorithm. About 63 per cent of galaxies that are in 2D groups with 7 or more members could also be considered to be in 3D groups. However, only about 45 per cent of 2D groups with 7 or more members would also be considered a 3D group with 7 or more members, rising to 80 per cent for 3D groups with 4 or more members. In general, then, we concluded that if our 2D groups trace 3D groups, then they must trace quite small 3D groups.

Chapter 4 subsequently sought the effect of the statistical lensing of 2QZ QSOs by foreground structure detected in the 2dF Galaxy Redshift Survey (2dFGRS) group sample (Eke et al. 2003). The 2dFGRS group sample is optimised to identify groups that are representative of underlying dark matter haloes, so should be well matched to our lensing models. 2QZ QSOs were cross-correlated against 2dFGRS groups with 5 or more members and we concluded that there is a significant anti-correlation between 2dFGRS groups and

2QZ QSOs but only on small scales; $\omega_{cq}(\theta < 2') = -0.16 \pm 0.05$. There is no significant anti-correlation on larger scales. We demonstrated that this result is not unique to the subsample of 2dFGRS groups with 5 or more members. A posteriori, we found that only the cross-correlation between a subset of 2dFGRS groups of $\sim 15$ or more members and 2QZ QSOs showed as strong an anti-correlation as that found in Chapter 3, and such groups are scarce enough that even that signal was at the limit of the statistics that could be probed by our data. By comparing the mass estimated for each group by Eke et al. (2003) to masses estimated by fitting the NFW lensing model described in Chapter 3, we concluded, unlike in Chapter 3, that there is no discrepancy between groups masses estimated from the statistical lensing of QSOs and group masses estimated in comparison to ΛCDM simulations.

Chapter 5 studied the cross-correlation of high-redshift ($z > 0.4$) 2QZ QSOs and $b_J < 20.5$ galaxies. We found that QSOs are anti-correlated with galaxies to the extent $\omega_{qg}(\theta < 10') = -0.007$ at nearly $3\sigma$ significance. There is no significant correlation between 2QZ stars and galaxies, suggesting that the QSO-galaxy anti-correlation is not some systematic of the 2dF observational technique. We measured the colours of 2QZ QSOs as a function of their distance from galaxies, as compared to a control sample of the same spatial distribution of QSOs but with colours selected at random from the 2QZ QSO distribution of colours. This analysis of colours suggested that absorption by dust in galaxies could account for at most 30 per cent of the QSO-galaxy anti-correlation, at the 95 per cent significance level. We demonstrated that, should dust in galaxies have a reddening law like that found by Calzetti, Kinney & Storchi-Bergmann (1994), dust could account for only a further 7-8 per cent of the anti-correlation.

We demonstrated that the anti-correlation between QSOs and galaxies is consistent with some effects that are expected if it arises due to the statistical lensing of QSOs. Notably, the anti-correlation evolves into a positive correlation as brighter samples of QSOs are cross-correlated against galaxies - although this effect is not significant. Somewhat unexpectedly, there is a stronger anti-correlation between radio-loud 2QZ QSOs and galaxy groups. This signal may be due to the effects of double magnification bias (Borgeest, von Linde & Refsdal 1991) but if so, could only arise due to some complicated form of this bias (Wyithe et al. 2003). If the QSO-galaxy anti-correlation were entirely due to the statistical lensing of QSOs, then galaxies would have to be highly anti-biased ($b \sim 0.05$) on scales of $0.1\,h^{-1}$Mpc. This is equivalent to saying that mass is highly clustered on small scales, as compared to galaxies, inconsistent with the ΛCDM paradigm. Our results are

particularly interesting when compared to the analysis of Gaztañaga (2003) who finds that galaxies are similarly anti-biased ($b \sim 0.1$) on small scales ($0.2\,h^{-1}$Mpc) by cross-correlating a bright SDSS QSO sample with SDSS galaxies and measuring a positive correlation between galaxies and QSOs. It is difficult to imagine a systematic effect that could cause both an anti-correlation and positive correlation between QSOs and galaxies, except for the statistical lensing of QSOs.

Some results in this thesis are at odds with the standard ΛCDM paradigm. Many of the discrepancies between this thesis and ΛCDM could be resolved if the integrated QSO number-count slope flattened considerably at magnitudes fainter than the limit of the 2QZ. In Chapter 6, we discussed two surveys we carried out to count QSOs in the magnitude range $20.5 \leq b_J < 22.5$. Our surveys, performed with WYFFOS (Corradi et al. 2001) and 2dF (Bailey & Glazebrook 1999), utilised the same UVX selection process used to target QSO candidates for the 2QZ, but our surveys used SDSS photometry. We transformed the magnitudes of star-like SDSS Early Data Release sources to the UKST bands and estimated that the resulting $b_J$ magnitudes would be accurate to $\pm 0.1$ mag. We discussed the observational technique and data reduction of our surveys at some length.

Ultimately, we identified 160 QSOs fainter than $b_J = 20.5$ that are *not* identified in the catalogue of Véron-Cetty & Véron (2001). We noted and demonstrated that WYFFOS suffers considerable light loss down fibres placed more than 20 arcminutes radius from the plate-centre (Corradi et al. 2001). The typical observational completeness of our two surveys (only considering the central 40 arcminutes of a WYFFOS field) was found to be 30 per cent. The typical spectroscopic completeness was found to be 40 per cent. After correcting our QSO numbers for incompleteness (and for dust absorption in our Galaxy), we found that, in the magnitude range $21 \leq b_J < 22$, the integrated QSO number-counts have a slope of $0.29 \pm 0.03$, highly consistent with the value inferred in Chapter 2 and used throughout this thesis. Finally, we discussed the 2dFSDSS collaboration's intention to extend the 2dF survey to fainter magnitudes. Using our faint QSO data, we estimated that $\sim 85$ per cent of $g < 21.15$ candidates targeted by the 2dFSDSS will be QSOs and $\sim 75$ per cent of $g \geq 21.15$ candidates targeted by the 2dFSDSS will be QSOs, higher than the QSO fraction identified in prior UVX surveys such as the 2QZ ($\sim 60$ per cent of which are QSOs).

## 7.3 Final Conclusions and Future Prospects

This thesis has dealt with the statistical lensing of QSOs and its implications for cosmological parameters. Many of the questions we have posed remain unanswered or incomplete. Though we have shown that that the distribution of high-density peaks in the galaxy distribution significantly influences the distribution of (redshift 0.4 and greater) 2QZ QSOs, and concluded that there may be more mass in groups than accounted for in the ΛCDM paradigm, we have applied quite simple modelling to the effect. High-resolution ray-tracing simulations of the effect of galaxy groups on QSOs would be beneficial. Large simulations of weak lensing on a cosmological scale are still in their infancy but do suggest greater magnification in ΛCDM universes than might be naively expected (Barber et al. 2000). However, to compare lensing through dark matter in simulations to the observed distribution of galaxies and groups will require a better understanding of how galaxies are biased relative to dark matter.

An alternative approach, studying 2dFGRS groups that are selected to mimic the distribution of NFW haloes in a ΛCDM cosmology, suggests no discrepancy between expected halo masses and lensing masses measured from statistical lensing. What this is telling us about the mass in our Universe where the density of the $b_J < 20.5$ galaxy distribution peaks on the plane of the sky is unclear. If ΛCDM correctly accounts for the distribution of mass in our Universe, then the anti-correlation between QSOs and galaxy groups selected as over-densities on the plane of the sky must arise due to some quirk of lensing near points where the projected number of $b_J < 20.5$ galaxies are particularly dense. To investigate this question will require extensive ray-tracing simulations of lensing magnification out to QSO redshifts coupled with a realistic model of galaxy bias, and might also need a survey of galaxy redshifts to $b_J < 20.5$, to investigate if there is a reason why QSOs seem to evade galaxy groups defined to this magnitude limit.

Coupled with the work of Gaztañaga (2003), we have found evidence that there is a consistent lensing signal in two independent samples of QSOs, samples where statistical lensing would boldly predict lensing signals quite different in form. This consistent lensing signal suggests that dark matter is highly anti-biased on small scales relative to galaxies. Again, the modelling used might be too simplistic. Galaxy bias may be very complicated on small scales. It is likely that much of theroretical cosmology in the next decade will study the complicated relationship between the light from baryons in the Universe and the dark matter component on small scales. Hopefully, alongside these theoretical

developments, better constraints on galaxy bias across a range of scales can be determined from lensing predictions based on a comparison of QSO and galaxy clustering in the completed Sloan Digital Sky Survey.

As the work outlined in this thesis has progressed, other authors have been using a second aspect of statistical gravitational lensing, namely, the shearing of background galaxy images by the differential gravitational potential across foreground structure, to determine cosmological parameters. These studies (Hoekstra et al. 2001, Hoekstra et al. 2002) yield values for $\Omega_m$ and the galaxy bias parameter, $b$, that are entirely consistent with the ΛCDM paradigm. Why there is a discrepancy between some aspects of this thesis and these galaxy shear measurements is not clear. The most obvious assumption that could influence measurements of QSO magnification but not galaxy shear, is the form of the faint-end of the integrated QSO number-counts. However, the surveys presented in Chapter 6 of this thesis suggest that the slope of the QSO number-counts fainter than the 2QZ limit is consistent with what would be extrapolated from the 2QZ data. Should it gain enough telescope time, the 2dFSDSS collaborative survey will definitively measure the slope of the QSO number-magnitude counts fainter than the 2QZ limit. An alternative reason for false measurements of QSO magnification is dust in foreground structure obscuring QSOs. Though we have shown in this thesis that this would need to be particularly grey dust, it might be interesting to survey QSOs by means other than their flux to determine dust-independent QSO numbers near galaxies or groups. Practically, this means that QSOs would have to be selected by their variability, or their lack of proper motion (Hartwick & Schade 1990). The largest QSO surveys in the time domain are currently performed on far smaller scales than the 2QZ (Rengstorf et al. 2001).

In the wake of constraints on the cosmological parameters provided by WMAP CMB data (Spergel et al. 2003) and looking forward to the even better constraints expected from PLANCK data (Vielva et al. 2003), many cosmologists would argue that the large-scale cosmology of our Universe is now well-known. We are entering an era of precision cosmology. The geometrical part of the ΛCDM paradigm has been upheld. Clustering results from large-scale structure affirm this geometry (Percival et al. 2001, Outram et al. 2003). We expect the next decade of cosmology to centre on the form of our Universe on smaller scales, on theories of how baryons coalesce and cluster relative to dark matter in our Universe, and on the general effect of gravitational lensing in our Universe on the brightness of sources we use to infer cosmological results (Perlmutter et al. 1999, Barber et al. 2000). Is the dark component of our Universe well represented by haloes of

collisonless Cold Dark Matter (Moore et al. 1999, Ghigna et al. 2000, Sand, Tommaso & Ellis 2002, Power et al. 2003)? Why can't we detect this dark matter (Paling 2003)? How do the baryons, which we can observe, trace the dark matter, which we cannot and how are these baryons redistributed on galactic scales by winds and supernovae (Adelberger et al. 2003)? Does lensing have an important effect on the general clustering statistics we use to measure the geometry of our Universe (Barber & Taylor 2003, Okamoto & Hu 2003)? We hope that effects discussed in this thesis, if further confirmed in larger QSO surveys such as the Sloan Digital Sky Survey and 2dFSDSS Survey, will continue to pose interesting observational questions to be considered by the cosmological theories developed over the next few decades. Early Twentieth Century physics taught us that a few small discrepancies at a time when a scientific discipline is considered essentially solved can trigger the scientific advances of the next 100 years. As such, posing new questions and highlighting discrepancies between observations and established theory should be considered the essential pursuit of the Twenty-First Century observational cosmologist.

# Appendix

# *Appendix A - Classical Cosmology*

In this appendix, we derive descriptions of the basic cosmological parameters used in this thesis but not defined elsewhere within it.

## A.1 The Cosmological World Model

The simplest metric line element for an expanding universe that is isotropic and homogenous on sufficiently large scales and consistent with the Theory of General Relativity (Einstein 1916) is given by

$$ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{A.1}$$

where $k$ is a constant that defines the geometry of the universal model for all time. For $k = 0$, the model reduces to a spherically symmetric space - called a "flat" universe. The case $k > 0$ represents a closed geometry and the case $k < 0$ an open geometry. The variables $r, \theta$ and $\phi$ denote spherical polar coordinates that flow with the expansion of the universe ("comoving coordinates"); $t$ is the cosmic time coordinate and $ds$ is the Lorentz invariant metric interval. Note that under the assumption of Universal homogeneity, $t$ is equivalent to the proper time, as different observes may agree on Universal parameters at the same time. The function $a(t)$ paramaterises the expansion factor of our Universe.

Equation A.1 is named the Robertson-Walker Metric after the authors who first realised its form (Robertson 1935, Walker 1936). Friedmann (1923) and Lemaitre (1927) had previously independently derived the form of Einstein's Field Equations of General Relativity (which relate the geometry of the universe to its content) in an expanding universe

$$\frac{\ddot{a}}{a} = \frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \tag{A.2}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8}{3}\pi G\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \tag{A.3}$$

where $p$ is the pressure due to, and $\rho$ is the density of, matter and radiation in a universe, and $G$ is the gravitational constant. The dot denotes a derivative with respect to cosmic time $t$. The cosmological constant, $\Lambda$, was added to these equations independently by Einstein (arising as a constant of integration in solving his field equations) in an attempt to provide for a non-expanding model of our Universe.

Hubble (1929) showed that our Universe *is* expanding and proposed the linear expansion law relating the velocities of galaxies $v$ and their distances $r$

$$v = H_o r \tag{A.4}$$

where $H_o$ is Hubble's Constant. Subsequently, the Robertson-Walker Metric (Equation A.1) and Friedmann Equations (Equations A.2 and A.3) gradually became the accepted fundamental cosmological equations of our era. In fact, Hubble's Law is a criterion of any expanding, isotropic, homogenous model.

It is simple to see there is a fundamental relationship between Hubble's constant as a function of cosmic time and the scale factor of Equations A.1, A.2 and A.3. The Robertson-Walker Metric (Equation A.1) shows that the proper distance at any particular cosmic time is some function of distance multiplied by the scale factor $r_{pr} = a(t)f(r)$, so the proper velocity is

$$v_{pr} = \dot{a}f(r) = \frac{\dot{a}}{a} r_{pr} \tag{A.5}$$

So, by analogy with Equation A.4, $H(t) = \dot{a}/a$. The Hubble Constant as a function of time is usually called the Hubble Parameter, to distinguish it from the small-scale Hubble's Constant (denoted $H_o$).

Consider the Hubble parameter, yielded by $\dot{a}/a$ in Equation A.3, as made up of separate contributions from matter in our Universe and from the Cosmological Constant. We define the critical density of a universe as the matter density in a flat ($k = 0$) universe with $\Lambda = 0$

$$\rho_{crit} = \frac{3H^2}{8\pi G} \tag{A.6}$$

allowing us to define a quantity that describes the flatness of the Universe

$$\Omega_{\rm m} = \frac{\rho}{\rho_{crit}} = \frac{8\pi G}{3H^2}\rho \tag{A.7}$$

Similarly, ignoring the matter contribution by setting $\rho = 0$ in Equation A.3, we find the contribution to the curvature of the universe from the Cosmological Constant is

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H^2} \tag{A.8}$$

Under these definitions, the Second Friedmann Equation becomes

$$\Omega_\mathrm{m} + \Omega_\Lambda = 1 + \frac{kc^2}{H^2 a^2} \tag{A.9}$$

Equation A.9 is a fair distillation of Einstein's Field Equations in the cosmological domain. The terms on the left hand side of Equation A.9 are constructed to contain the matter content of the Universe, where the terms on the right describe the geometry of a universe. Sometimes, the term $\frac{kc^2}{H^2a^2}$ is denoted $\Omega_k$, the "critical curvature" of our Universe. The significance of $k$ is now apparent. If $k = 0$, the case of a "geometrically flat" universe, then the critical matter terms sum to 1. In a $\Lambda = 0$, flat universe, $\Omega_\mathrm{m} = 1$ and the density of the universe is critical. For an open ($k < 0$) cosmology, the $\Omega$ terms will total more than 1. For a closed ($k > 0$) geometry, the $\Omega$ terms will total less than 1.

Recent measurements of the angular power spectrum of the Cosmic Microwave Background suggest $\Omega_\mathrm{m}h^2 = 0.14 \pm 0.02$ and $h = 0.72 \pm 0.05$ (Spergel et al. 2003). Measurements assuming that supernovae are standard candles (Perlmutter et al. 1999) suggest $0.8\Omega_\mathrm{m} - 0.6\Omega_\Lambda \sim -0.2 \pm 0.1$. Key Project measurements of Hubble's Constant (Freedman et al. 2001) suggest $h = 0.72 \pm 0.08$ ($100h\,\mathrm{km\,s^{-1}\,Mpc} = H_o$). Measurements of the statistical distribution of large scale structure (Percival et al. 2001, Dodelson et al. 2002, Outram et al. 2003) are converging at around $\Omega_\mathrm{m}h = 0.2 \pm 0.05$. Probably any two of these observations (consider Efstathiou et al. 2002) and certainly any three of them are sufficient to say with some confidence that $\Omega_\mathrm{m} = 0.3, \Omega_\Lambda = 0.7$ and k = 0.

## A.2 The Redshift of Photons

The Robertson-Walker Metric is readily used to demonstrate that photons that are emitted at one time in an expanding cosmology and received at another time will redden. Consider a photon emitted at time $t_e$, which travels a distance $r$ to be received by an observer at the present time $t_o$. Photons travel along null geodesics, so we may set $ds = d\theta = d\phi = 0$ in Equation A.1

$$\int_{t_e}^{t_o} \frac{cdt}{a(t)} = \int_0^r \frac{dr}{\sqrt{1 - kr^2}} = f(r) \tag{A.10}$$

Where $f(r)$ simply denotes a function of $r$. Light emitted a small time later, $t_e + \delta t_e$, will be received a small time later, $t_o + \delta t_o$, so

$$f(r) = \int_{t_e+\delta t_e}^{t_o+\delta t_o} \frac{cdt}{a(t)} = \int_{t_e}^{t_o} \frac{cdt}{a(t)} - \int_{t_e}^{t_e+\delta t_e} \frac{cdt}{a(t)} + \int_{t_o}^{t_o+\delta t_o} \frac{cdt}{a(t)} \quad \text{(A.11)}$$

$$= f(r) - \frac{c\delta t_e}{a(t_e)} + \frac{c\delta t_o}{a(t_o)} \quad \text{(A.12)}$$

This reduces to

$$\frac{c\delta t_o}{a_o} = \frac{c\delta t_e}{a} \quad \text{(A.13)}$$

where $a(t_o)$ has been shortened to $a_o$ and $a(t_e)$ to $a$. In the previous equation, $c\delta t$ could be identified with the wavelength ($\lambda$) of the photon, meaning

$$\frac{\lambda_o}{\lambda_e} = \frac{a_o}{a} \quad \text{(A.14)}$$

If our Universe is expanding, then $a_o > a$, so $\lambda_o > \lambda_e$. The photon received at the present time is redder than when it was emitted. The cosmological redshift, $z$ is defined as the fractional reddening in a photon's wavelength between emission and reception

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{a_o - a}{a}. \quad \text{(A.15)}$$

## A.3 Hubble's Parameter as a Function of Redshift

The Second Friedmann Equation (A.3) may be combined with Equation A.15 to determine the Hubble Parameter in terms of measurable quantities. The present value of $\Omega_{\mathrm{m}}$ is given by Equations A.7 and A.6 with the Hubble Parameter set to its present value (Hubble's Constant, $H_o$)

$$\Omega_{\mathrm{m,o}} = \frac{\rho_o}{\rho_{crit,o}} = \frac{8\pi G}{3H_o^2}\rho_o \quad \text{(A.16)}$$

The present value of $\Omega_\Lambda$ is similarly yielded via Equation A.8

$$\Omega_{\Lambda,o} = \frac{\Lambda c^2}{3H_o^2} \quad \text{(A.17)}$$

In the matter-dominated era, the density of part of the Universe is always inversely proportional to the volume of that part, meaning, $\rho \propto a^{-3}$, where $a$ is the scale factor at the time the density is measured. This means $\rho = \left(\frac{a_o}{a}\right)^3 \rho_o$, which is to say $\rho =$

$(1+z)^3\rho_o$, given the definition of the redshift, $z$, in Equation A.15. The Second Friedmann Equation (A.3) can then be written

$$H(t)^2 = \frac{8}{3}\pi G(1+z)^3\rho_o - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \tag{A.18}$$

Substituting from Equations A.17 and A.16 for the present-day, measurable values of $\Omega_{\mathrm{m,o}}$ and $\Omega_\Lambda$, we find

$$H(t)^2 = H_o^2\Omega_{\mathrm{m,o}}(1+z)^3 - \frac{kc^2}{a^2} + H_o^2\Omega_{\Lambda,o} \tag{A.19}$$

To determine the present day value of $k$, which is, by definition, a constant for all times and places, we consider Equation A.19 with all terms measured at the present day, so $H(t) = H_o, a = a_o$ and $z = 0$

$$H_o^2 = H_o^2\Omega_{\mathrm{m,o}} - \frac{kc^2}{a_o^2} + H_o^2\Omega_{\Lambda,o} \tag{A.20}$$

so

$$-kc^2 = (1 - \Omega_{\mathrm{m,o}} - \Omega_{\Lambda,o})a_o^2H^2 = (1 - \Omega_0)a_o^2H^2 \tag{A.21}$$

Substituting this expression back into Equation A.19 and once more using Equation A.15 for the redshift, we find

$$H(z)^2 = H_o^2\left[(1 - \Omega_0)(1+z)^2 + \Omega_{\mathrm{m,o}}(1+z)^3 + \Omega_{\Lambda,o}\right] \tag{A.22}$$

which allows the Hubble Parameter to be determined for any redshift in terms of the present day measurables. When using this equation, we shall usually drop the cumbersome 0 = "present-day"subscript on $\Omega_{\mathrm{m}}$ and $\Omega_\Lambda$.

## A.4 Cosmological Distances

It is straightforward to show how redshift, $z$, evolves with cosmic time, $t$. We use the fact that $\frac{da}{dt} \equiv \dot{a}$ can be identified as $\dot{a} = aH$ using the definition of the Hubble Parameter (see Equation A.5) and differentiate the definiton of redshift (Equation A.15) with respect to the scale parameter, $a$, to see

$$\frac{dz}{dt} = \frac{dz}{da}\frac{da}{dt} = H(1+z) \tag{A.23}$$

We have previously shown the path of a radially moving photon in Equation A.10

$$\frac{cdt}{a} = \frac{dr}{\sqrt{1-kr^2}} \tag{A.24}$$

We will denote the radial comoving distance traced by such a photon from an object of interest in the Universe by $\chi$. As the comoving coordinates of objects in the Universe do not change with time (moving, as they do, with the expansion), the scale factor is the same for all times (the present scale is equivalent to the scale at any time). Setting dt $= d\phi = d\theta = 0$ and $a = a_o$ in the Robertson-Walker Metric (Equation A.1, the radial comoving distance element is given by

$$d\chi = a_o \frac{dr}{\sqrt{1-kr^2}} \tag{A.25}$$

which we can equate with Equation A.24 and the definition of redshift to yield

$$\frac{d\chi}{dt} = \frac{a_o}{a}c = (1+z)c \tag{A.26}$$

We combine Equations A.23 and A.26 to give

$$\frac{d\chi}{dz} = \frac{d\chi}{dt}\frac{dt}{dz} = \frac{c}{H} \tag{A.27}$$

an equation that may be integrated to give the radial comoving distance to any object in our Universe, provided its redshift is known. In this thesis, we mainly use Equation A.23 and Equation A.27, together with the equation for Hubble's Parameter in terms of redshift (Equation A.22) to determine distances from redshift and values of the present day cosmological parameters $(c, H_o, \Omega_m, \Omega_\Lambda)$.

Although the most readily conceptual cosmological distance, the radial comoving distance is not directly measurable. Expressions in gravitational lensing models often simplify greatly when related to a distance concept that can, in principle, be directly measured - *angular diameter distance*. In terms of measurements, the angular diameter distance is the distance at which a standard rod of certain proper length subtends a given angle. In terms of the cosmological world model, the angular diameter distance is defined similarly to the comoving distance but because galaxies will appear smaller as they recede with the expansion, the angular diameter distance moves with the scale factor, $a$. It turns out that the physically meaningful model of the angular diameter distance, $D_A$ at a given redshift, $z$ out to the Robertson-Walker metric radial coordinate $r$ is

$$D_A(z) = a(z)r(z) = \frac{a_o r(z)}{(1+z)} \tag{A.28}$$

Given Equations A.25 and A.27 we find

$$\frac{dr}{dz} = \frac{dr}{d\chi}\frac{d\chi}{dz} = \frac{\sqrt{1-kr^2}}{a_o}\frac{c}{H} \tag{A.29}$$

which can be integrated to give $r(z)$, which in turn yields $D_A(z)$. It is possible but not straightforward to generally determine the angular diameter distance between two redshifts, such as might be necessary to characterise the distance between a gravitational lens and the source it is lensing, by numerically integrating a second-order differential equation known as the Dyer-Roeder Equation (Dyer & Roeder 1974, Dyer & Roeder 1981), as outlined in Section 5.1 of Schneider, Ehlers & Falco (1992). The Dyer-Roeder Equation can also model the gross effect of gravitational lensing through the field of in an inhomogenous cosmology on the distance determinations. However, the Dyer-Roeder Equation can quickly become complicated in inhomogenous universes or universes with a Cosmological Constant (Demianski et al. 2000, Kantowski 2000). In this thesis, we tend to use the practical approach of Kayser, Helbig & Schramm (1997) when calculating the angular diameter distance between two redshifts. Unless otherwise mentioned, all angular diameter distances in this thesis are calculated assuming a smooth (homogenous) universe.

# Appendix

# *Appendix B - Modern Cosmology*

In this appendix, we discuss the main statistics used to test modern cosmology, namely the two-point correlation function and its Fourier-space twin the power spectrum. We also discuss how galaxy bias is measured using either statistic.

## B.1 The Two-point Correlation Function

The two-point correlation function, $\xi(r)$ is a measure of how clustered a set of points in space are (Peebles 1980). Consider the following measure of the probability ($\mathrm{d}P$) of finding pairs of sources of mean number density $\bar{n}$ within volume $\mathrm{d}V$ separated by distance $r$

$$\mathrm{d}P_{12} = \bar{n}_1 \bar{n}_2 \left[1 + \xi(r)\right] \mathrm{d}V_1 \mathrm{d}V_2 \tag{B.1}$$

Clearly, if $\xi(r) = 0$, then we obtain the expected expression for two sets of points that are randomly distributed relative to each other

$$\mathrm{d}P_{R_1 R_2}(r) = \bar{n}_{R_1} \bar{n}_{R_2} \mathrm{d}V_{R_1} \mathrm{d}V_{R_2} \tag{B.2}$$

but, if $\xi(r) > 0$, or $\xi(r) < 0$ then we have a greater or lesser chance of finding a point at a given distance from a second point and our two sets of points either cluster around each other or avoid each other, as might be expected for real data

$$\mathrm{d}P_{D_1 D_2}(r) = \bar{n}_{D_1} \bar{n}_{D_2} \left[1 + \xi(r)\right] \mathrm{d}V_{D_1} \mathrm{d}V_{D_2} \tag{B.3}$$

If we wish to measure $\xi(r)$, then we need only compare the distribution of data-points to a set of random points. Taking Equations B.2 and B.3 in ratio

$$\xi(r) = \frac{\mathrm{d}P_{D_1 D_2}(r)}{\mathrm{d}P_{R_1 R_2}(r)} \frac{\bar{n}_{R_1}}{\bar{n}_{D_1}} \frac{\bar{n}_{R_2}}{\bar{n}_{D_2}} \frac{\mathrm{d}V_{R_1} \mathrm{d}V_{R_2}}{\mathrm{d}V_{D_1} \mathrm{d}V_{D_2}} - 1 \tag{B.4}$$

In practice, one usually integrates over the volumes and probabilities in these expressions, simply counting the number of data-data ($D_1D_2$) pairs and random-random ($R_1R_2$) pairs in a given volume of sky. With a little thought, it is possible to construct a number of estimators of $\xi(r)$. A simple estimator that is often used, and that is accurate when the survey volume is large enough that the number density of sources is well-known (i.e. a large enough part of the Universe is sampled that clustering at the edges of the field do not have any bearing on the estimation of the number density of sources), is

$$\xi(r) = \frac{D_1D_2(r)}{D_1R_2(r)}\frac{n_{R_2}}{n_{D_2}} - 1 \tag{B.5}$$

This estimator is rapid, as a set of random points need only be constructed to mimic one of the sets of data points.

If we are interested in working in the plane of the sky rather than over the entire volume of a survey (perhaps as redshifts of sources aren't available), then one can define an analogous *angular two-point correlation function*, $\omega(\theta)$ simply by making the substitutions $\xi \mapsto \omega$, $r \mapsto \theta$, $\mathrm{d}V \mapsto \mathrm{d}\Omega$. If we are interested in measuring the intrinsic clustering of a particular set of objects, rather than the relative clustering of two sets of objects, we can define the *auto-correlation* function simply by dropping any subscripts in the above equations, so that we are only dealing with one species.

## B.2 The Power Spectrum

Though we never directly measure the power spectrum in this thesis, it is frequently the statistic of choice in modern studies of clustering, and we often discuss it in the body of the text, so we will briefly outline its meaning here.

Consider Equation B.5 but written to express the *auto-correlation* function (i.e. $D_1 = D_2$ and $R_1 = R_2$). Assuming we are sitting at a given coordinate in space and counting the number of data points and random points at a distance $r$ away, and further rewriting the number of data points (at a distance, $r$) divided by the mean number of data points (at a distance, $r$) in terms of the relative *densities*, $\rho$, of data points, we find

$$\delta(r) = \frac{\rho(r)}{\bar{\rho}} - 1 = \frac{\rho(r) - \bar{\rho}}{\bar{\rho}} \tag{B.6}$$

where the number of random points divided by the mean number of random points at

any scale is just the expected density at any coordinate. The quantity $\delta(r)$ is called the density contrast, and characterises fluctuations in the density field at any coordinate.

We have made a subtle assumption here. As we constructed the two-point auto-correlation so that it measured the mean density of points at any distance from a given point, it seems that at any scale $r$, we should recover the exact same value of the density of points. This is not necessarily the case, we can construct situations where the distribution of points is highly skewed. In an extreme case, if one-half of the points in a particular density field are extremely clustered but the other half avoid each other to the exact same degree, then we will measure zero for the correlation function, as on average, the points are unclustered. However, this skewed situation is very different from a distribution of randomly placed points, which are genuinely unclustered. A density field where the mean number of points around any given point is an exact representation of the distribution of points on that scale is called a *Gaussian Random Field.* Clearly, a Gaussian Random Field, where over any scale, the same (mean) number of points is repeated, is periodic. Under the assumption that Equation B.6 is describing a Gaussian Random Field, the density contrast is a periodic function and can thus be Fourier expanded

$$\delta(r) = \frac{1}{8\pi^3} \int \delta_k e^{-ik.r} d^3k \tag{B.7}$$

for modes in the range $0 \rightarrow 2\pi$ in each spatial dimension. The amplitude of the modes of the density field are given by the Fourier transform of this equation

$$\delta_k = \int \delta(r) e^{ik.r} d^3r \tag{B.8}$$

The power spectrum, $P(k)$ of a density field for any wavenumber $k$ is defined as the square modulus of the amplitude of the modes at that wavenumber

$$P(k) = |\delta_k|^2 \tag{B.9}$$

In practice, the calculation of the power spectrum for observational data is more complicated than outlined here. A good reference is Tadros & Efstathiou (1996). As might be expected from the route we have taken to describe the power spectrum, the power spectrum is the Fourier transform of the two-point auto-correlation function. Note that both the two-point auto-correlation function and the power spectrum contain enough

information to characterise a density field if *and only if* the field is a Gaussian Random Field. Fields with skewness must be described by higher-order statistics. Inflation predicts that the CMB should be Gaussian and Komatsu et al. (2003) have recently analysed WMAP data, suggesting the CMB is very close to Gaussian.

## B.3 Galaxy Bias

As discussed in Chapter 1, galaxies (or other luminous sources) in our Universe only represent a small fraction of the mass. In the CDM picture, huge gravity wells of dark matter agglomerate and merge over time, trapping baryonic matter in their thrall. The baryons, which, unlike the CDM, can collide and interact, can heat through shocks and nuclear reactions, creating the light we see in the Universe. Thus, the light traces the CDM but does not necessarily cluster in the same manner as the CDM. The function that describes how galaxies cluster relative to the underlying dark matter is the galaxy bias function. Representing the clustering of galaxies using the two-point auto-correlation function with subscript "g" and the clustering of mass using the two-point auto-correlation function with no subscript, we can write

$$\xi_g(r) = f(b)\xi(r) \tag{B.10}$$

where $f(b)$ represents some unknown bias function.

Many attempts have been made to model the form of the bias function (Baugh, Cole & Frenk 1996, Seljak 2000, Cooray & Sheth 2002) but the results can be quite different, depending on the assumptions made. In the simplest and probably most widely-used bias model (Kaiser 1984) the dark matter is assumed to be a Gaussian Random Field and galaxies to form in places where a fluctuation in the dark matter field peaks above a certain arbitrary threshold, $\nu$. This gives the following relation

$$\xi_g(r) = b^2\xi(r) \tag{B.11}$$

where $b$, the bias parameter, is the number of standard deviations above the threshold that the dark matter must fluctuate to trigger galaxy formation (i.e. $b = \nu/\sigma$ where $\sigma$ is the standard deviation of the Gaussian Random Field of the dark matter). In practice, on scales of a few Megaparsecs and above, most studies find $b \sim 1$ (Verde et al. 2002, Lahav

et al. 2002, Hoekstra et al. 2002), meaning galaxies form in moderate $1\sigma$ fluctuations and there is no particular bias. If $b < 1$ (i.e. the threshold necessary for galaxies to form is lower than $1\sigma$) then galaxies predominately form in regions where the mass is not particularly clustered (the CDM is far more clustered than the galaxies). If $b > 1$ then galaxies form in the most clustered regions of mass (the galaxies are more highly clustered than the CDM). This simple model of galaxy bias is often called "high peaks" bias. Currently, the auto-correlation function of underlying matter in a universe is inferred from simulations of dark matter, and the galaxy bias is estimated by a comparison to the observed auto-correlation function of galaxies (Baugh, Cole & Frenk 1996).

# Appendix

# *Appendix C - Gravitational Lensing*

In this appendix, we discuss some of the gravitational lensing theory used to construct models in this thesis.

## C.1 The Lens Equation

General Relativity predicts that mass will lens light, deflecting light rays from the source in the lensing plane by the deflection angle, $\alpha$

$$\alpha = \frac{4GM(<\xi)}{\xi c^2} \tag{C.1}$$

(in the Schwarzschild metric) for a mass $M$ at impact parameter $\xi$, which, turns out to be exactly twice the prediction of Newtonian Theory.

In Fig. C.1, we plot the lensing geometry in the case of a circular-symmetric lens with radius much smaller than the impact parameter (i.e. space described by the Schwarzschild Metric). Light rays are emitted by a source and take one of two routes around a lensing mass. We only specifically consider one route, as the geometry is equivalent for either route. In the cases discussed in this thesis (and in nearly all cosmological situations) $\theta, \theta_q, \alpha << 1$ (i.e. $\xi << 4GM/c^2$, which is twice the Schwarzschild Radius of the lens), so the distance $\eta$ (marked in Fig. C.1) may be expressed

$$\eta = \theta_q D_s = \theta D_s - \alpha D_{ls} \tag{C.2}$$

In the one-dimensional case plotted in Fig. C.1, the Lens Equation is expressed

$$\theta_q = \theta - \frac{D_{ls}}{D_s}\alpha \left( = \theta - \frac{D_{ls}}{D_s}\frac{4GM(<\xi)}{\xi c^2} \right) \tag{C.3}$$

where we have bracketed the final part of the equation to highlight the fact that the deflection angle $\alpha$ is only given by Equation C.1 in the specific case where the coordinate

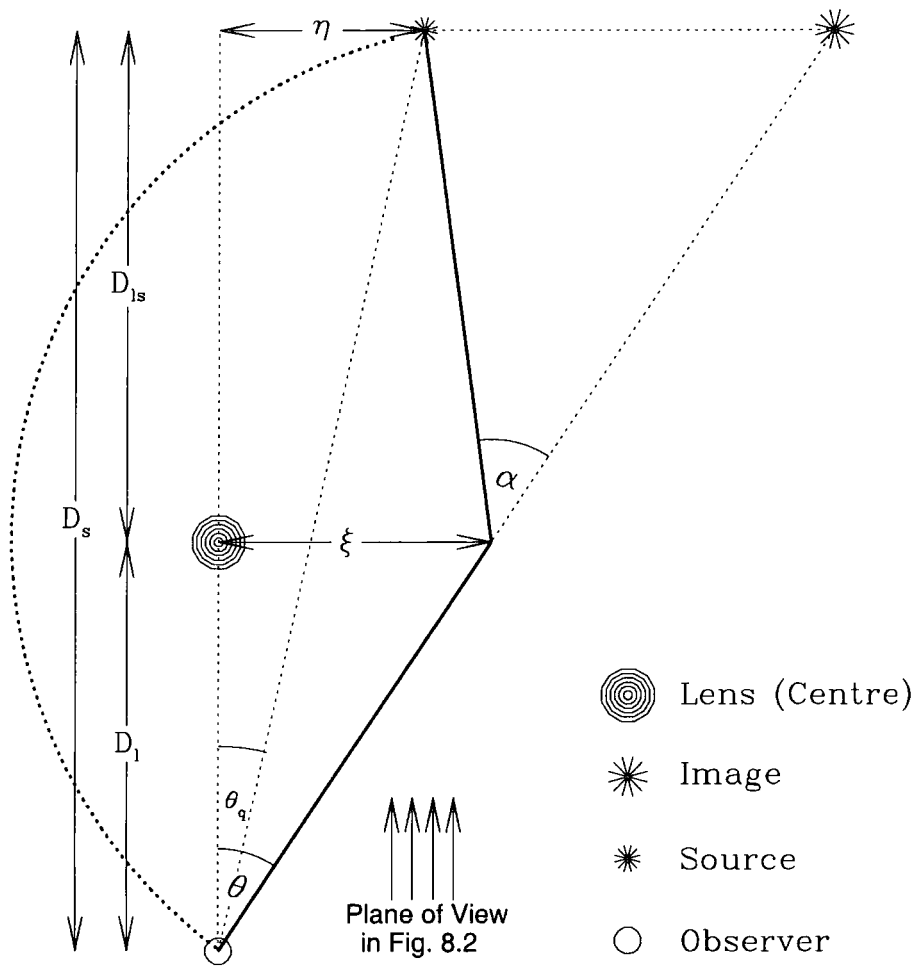


Figure C.1: Light from a source at angular diameter distance $D_s$ that is a distance $\eta$ from the optical axis (drawn here as the line adjoining the observer and the lens), would, in the absence of gravitational lensing, traverse the path at angle $\theta_q$ to the observer. However, a massive lens at distance $D_l$ can make the light ray deviate as it passes the lens at impact parameter $\xi$, bending the ray back towards the observer. The image then appears to have come from further than $\eta$ from the optical axis. The deviation angle, $\alpha$ is predicted by General Relativity. Note that a second light path exists on the other side of the lens (drawn as the curved dotted line), so, in actuality, there are two images, with similar geometry.

space is described by the Schwarzschild metric. Equation C.3 is easily extended to the two-dimensional case, if the various angles involved are considered as vectors.

Note that in the spherically symmetric one-dimensional case, where the two possible images of the source appear equidistant from the lens, the source must lie on the optical axis and $\eta = \theta_q = 0$, so

$$\theta = \frac{D_{ls}}{D_s}\frac{4GM}{\xi c^2} = \frac{D_{ls}}{D_s}\frac{4GM}{D_l \theta c^2} \tag{C.4}$$

and the two images form at the "Einstein Radius"

$$\theta_E^2 = \frac{4GM}{c^2}\frac{D_{ls}}{D_s D_l} \tag{C.5}$$

In the two-dimensional (circular-symmetric) case, these two images sweep out a continuous circle around the lens - the "Einstein Ring". If the symmetry were elliptical, the Einstein Ring would be an ellipse (and so forth). Using this definition of the Einstein Radius, we can rewrite Equation C.3 as

$$\theta_q = \theta - \frac{\theta_E^2}{\theta} \tag{C.6}$$

a quadratic with solutions

$$\frac{1}{2}\left(\theta_q \pm \sqrt{\theta_q^2 + 4\theta_E^2}\right) \tag{C.7}$$

If we write $\theta_q = n\theta_E$ these two solutions can be expressed as

$$\theta_{1,2} = \left(\frac{n}{2} \pm \sqrt{\left(\frac{n}{2}\right)^2 + 1}\right)\theta_E \tag{C.8}$$

and the difference between these two solutions is

$$\theta_1 - \theta_2 = \left(\sqrt{\left(\frac{n}{2}\right)^2 + 1}\right) 2\theta_E \tag{C.9}$$

meaning that the closest approach of the two images is twice the Einstein Radius, when $\theta_q = \eta = 0$. Hence, every image within the Einstein Radius has a pair on the opposite side of the lens that lies exterior to the Einstein Radius. Closer analysis of Equation C.8 shows that images at infinity map to the centre of the lens and, in fact, each point in the source plane has two images, one interior and one exterior to the Einstein Radius.

## C.2 Magnification Bias

In Fig. C.2 we plot the (infinitesimal) cross-section of rays of light that have been emitted from the source in Fig. C.1 and lensed into the eye of the observer. In fact, we are looking along the cluster of arrows labelled in Fig. C.1. We state without proof that *surface brightness is conserved during the process of gravitational lensing*, so an image of a source has the same surface brightness as the source itself (see, e.g., section 3.8 of Schneider, Ehlers & Falco 1992). As surface brightness (for a sufficiently small elemental area) is the product of flux and solid angle, a distortion in the cross-sectional area of a light-ray magnifies (or de-magnifies) the flux of an image as compared to its source. The magnification can then be expressed entirely as the ratio of the solid angles subtended by an image and a source. If the image has area $A_I$ and is at distance $D_l$ and the source has area $A_S$ and is at distance $D_s$, then the ratio of the emitted solid angle $\omega_e$ and observed solid angle $\omega_o$ is

$$\mu = \frac{\omega_e}{\omega_o} = \frac{A_I}{A_S}\left(\frac{D_s}{D_l}\right)^2 \tag{C.10}$$

where $\mu$ is the magnification.

Now, In Fig. C.2, we illustrate infinitesimal area elements subtended by the image (in the lens plane) and the source (projected into the lens plane). The geometry of the situation implies that $A_S = D_s^2\theta_q \mathrm{d}\theta_\mathrm{q} \tan\phi$ and $A_I = D_l^2\theta \mathrm{d}\theta \tan\phi$, which, together with Equation C.10 implies

$$\mu = \left|\frac{\theta}{\theta_\mathrm{q}}\frac{\mathrm{d}\theta}{\mathrm{d}\theta_\mathrm{q}}\right| \tag{C.11}$$

where the absolute value is taken only to highlight the fact that the second image (that appears on the opposite side of the lens) will, in vectorial notation, have a negative value ( by analogy with the theory of optics, it is a mirror-image of the first image).

One can use Equation C.8 and Equation C.11 to show that the two images in the one-dimensional circular symmetric case considered in Section C.1 will have magnifications

$$\mu_{1,2} = \frac{n^2+2}{2n\sqrt{n^2+4}} \pm \frac{1}{2} \tag{C.12}$$

which shows (remembering $\theta_q = n\theta_E$) that exactly on the optical axis, the images that form at the Einstein Radius are infinitely magnified. Further, when an image-pair forms

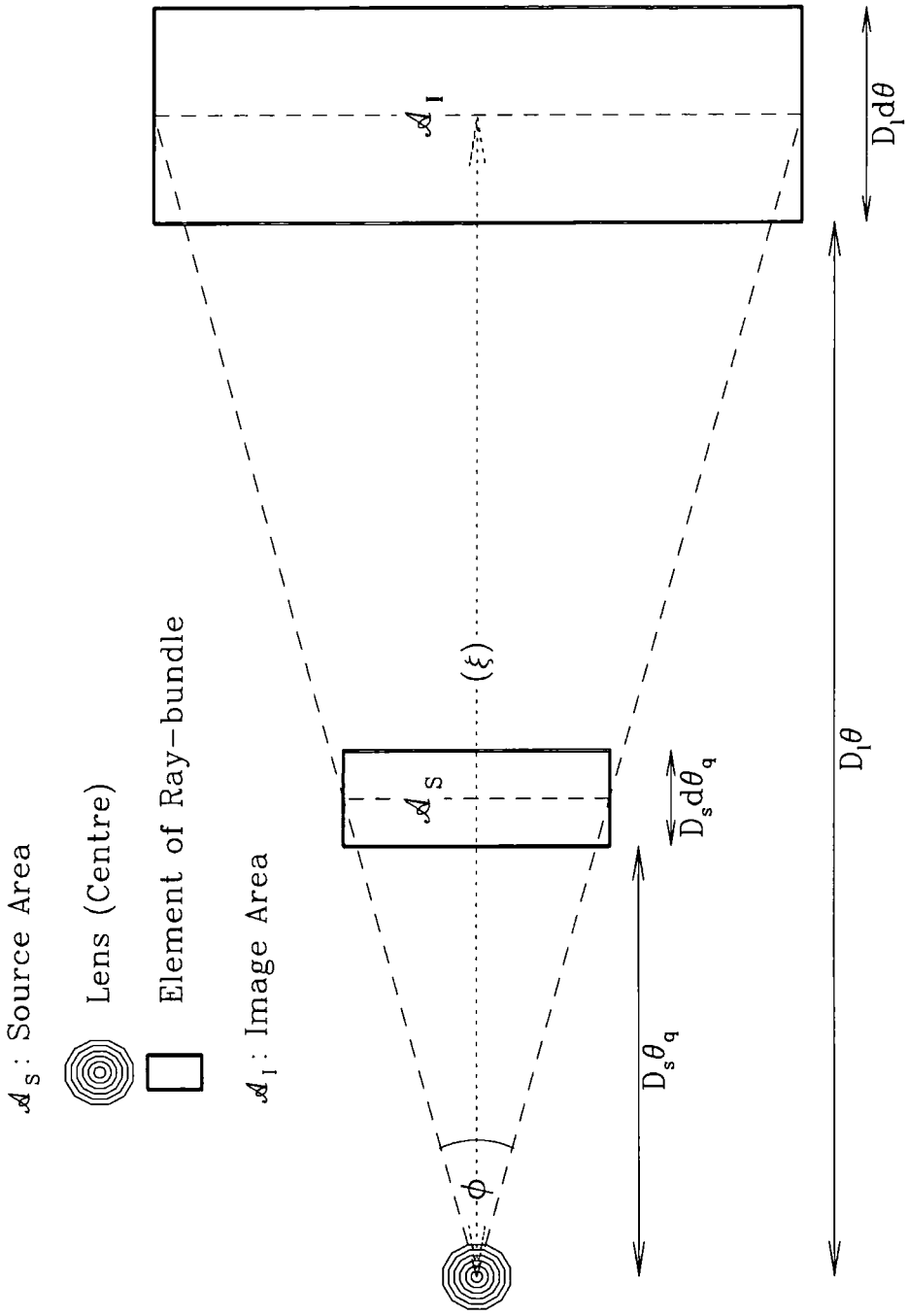


Figure C.2: Gravitational Lensing conserves surface brightness. For an infinitesimal source, surface brightness is the product of the source flux and solid angle. Thus if the solid angle subtended by a source is altered by lensing, the flux received can be larger or smaller than expected in the absence of lensing - this is called magnification. Above, we plot two elemental bundles of light rays. The source bundle is the projection into the lens plane of the emitting source in the absence of lensing. The image bundle is the image of the source bundle in the lens plane in the presence of lensing. Lensing alters the (infinitesimal) area of a ray bundle.

with one image at infinity and the other at the centre of the lens (see the last sentence of Section C.1), the image at infinity is the same magnitude as the source and the image at the centre of the lens is invisible. In general, one of the image pairs that form inside and outside the Einstein Ring is magnified and the other is de-magnified.

Let us consider this magnification over a large portion of sky. If we have surveyed a sample of objects to a limiting magnitude $m$, in the absence of lensing, the number of sources we will observed will be $N(< m)$. However, magnification by lensing will have two effects on our survey. First, as the magnification is the ratio of the lensed and unlensed flux of an object, some objects that were two faint to be included in our sample in the absence of lensing will be magnified into the sample $N(< m) \rightarrow N(< m + 2.5 \log \mu)$. Secondly, as magnification arises through a distortion that increases the area behind a lens, the area we are observing is smaller than that we would observe in the absence of lensing, meaning the number of objects observed is reduced by a factor $1/\mu$. Hence, the number of objects observed in the presence of lensing as compared to the number observed in the absence of lensing $q$ is given by (Narayan 1989)

$$q = \frac{1}{\mu} \frac{N(< m + 2.5 \log \mu)}{N(< m)} \tag{C.13}$$

## C.3 The General Lens

In two dimensions, the impact parameter becomes a vectorial quantity in the lens plane. In three dimensions, the propogation of a light ray through an extended mass distribution may be envisaged by projecting the entire mass distribution onto a single plane, then treating this as the lens plane (this is called the *thin lens approximation*). Compared to the impact parameter, $\vec{\xi}$, in the lens plane, the impact parameter in other planes will be $\vec{\xi} - \vec{\xi}'$, where $\vec{\xi}'$ accounts for the deviation of the light ray as it propogates through the extended lensing mass. The mass of the extended lens may be similarly thought of as projected onto the lens plane, meaning the surface mass density at any point on the lens plane $\Sigma(\vec{\xi})$ is the integral of the density encountered by the light ray along the density distribution of the lens, $\rho(\vec{r})$. If the vector through the lensing distribution is $\vec{r} = (\vec{\xi}, r)$ then

$$\Sigma(\vec{\xi}) = \int_0^{D_s} \rho(\vec{r}) \mathrm{d}r \tag{C.14}$$

By analogy with Equation C.1, the deflection angle through the entire lensing medium, being the sum of the mass encountered at each impact angle at each impact parameter along the propogation of the light ray, is given by

$$\vec{\alpha}(\vec{\xi}) = \frac{4G}{c^2}\sum_i \mathrm{dm_i}(\vec{\xi'}, \mathrm{r})\frac{\vec{\xi} - \vec{\xi'}}{|\vec{\xi} - \vec{\xi'}|^2} \tag{C.15}$$

where $\xi$ is now constructed as a vector rather than the scalar of Equation C.1. Projecting the mass distribution into the lens plane, and integrating over the entire area in the lens plane in the continuous limit, this is equivalent to

$$\vec{\alpha}(\vec{\xi}) = \frac{4G}{c^2}\int\int_0^{D_s} \rho(\vec{\xi'}, r)\mathrm{dr}\frac{\vec{\xi} - \vec{\xi'}}{|\vec{\xi} - \vec{\xi'}|^2}\mathrm{d}^2\xi' = \frac{4\mathrm{G}}{\mathrm{c}^2}\int \Sigma(\vec{\xi'})\frac{\vec{\xi} - \vec{\xi'}}{|\vec{\xi} - \vec{\xi'}|^2}\mathrm{d}^2\xi' \tag{C.16}$$

Note that the Lens Equation (Equation C.3) still holds, in vector form. In the one-dimensional case considered when introducing the Lens Equation, we discussed the Einstein Ring. By analogy with the Einstein Ring, in three dimensions, we consider a circle projected onto the lens plane that has constant surface mass density, so $\xi' = 0$ and $\Sigma(\vec{\xi'}) = \Sigma$. In this case, Equations C.16 and the Lens Equation combine to give

$$\vec{\theta_q} = \vec{\theta} - \frac{D_{ls}}{D_s}\frac{4G}{c^2}\frac{\pi\vec{\xi}^2}{|\xi|}\Sigma = \vec{\theta} - \frac{D_{ls}D_l}{D_s}\frac{4\pi G}{c^2}\vec{\theta}\Sigma \tag{C.17}$$

where we have reintroduced $\vec{\xi} = D_l\vec{\theta}$. In the on-axis case where $\theta_q = \eta = 0$ (by analogy with Fig. C.1), we define the critical surface mass density $\Sigma_{crit}$

$$\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}} \tag{C.18}$$

A comparison to Equation C.5 shows that the critical surface mass density is analogous to the total mass of a point-like lens being distributed around the Einstein Radius.

How close the surface mass density is to critical at any point is denoted by the dimensionless surface mass density, or convergence, $\kappa(\vec{\theta})$

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_l\vec{\theta})}{\Sigma_{crit}} \tag{C.19}$$

and a mass distribution that has $\kappa \geq 1$ somewhere will produce some multiple images, much as discussed in the simplified case of Section C.1 (see Schneider, Ehlers & Falco 1992; section 5.4).

Now, Equation C.16 can be rescaled using the convergence (and the substitution $\vec{\xi} = D_l\vec{\theta}$), becoming

$$\frac{D_{ls}}{D_s}\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi}\int \kappa(\vec{\theta'})\frac{\vec{\theta}-\vec{\theta'}}{|\vec{\theta}-\vec{\theta'}|^2}\mathrm{d}^2\theta' \tag{C.20}$$

and we can define the (scalar) lensing potential, $\Psi$, using

$$\frac{D_{ls}}{D_s}\vec{\alpha} = \nabla\Psi \;\; ; \;\; \Psi(\vec{\theta}) = \frac{1}{\pi}\int \kappa(\vec{\theta'})\ln|\vec{\theta}-\vec{\theta'}|\mathrm{d}^2\theta' \tag{C.21}$$

By analogy with Poisson's Equation, it can be shown that (Blandford & Narayan 1986)

$$\nabla.\left(\frac{D_s}{D_l}\vec{\alpha}\right) = \nabla^2\Psi = 2\kappa \tag{C.22}$$

which isn't necessarily obvious without seeing a derivation of Poisson's Equation (Binney & Tremaine 1987). Much of what remains will be similarly outlined rather than derived in much detail, as to detail it would grossly lengthen this already grossly lengthy appendix.

### C.3.1 Geometric Distortion of an Image

As discussed in Section C.2, gravitational lensing conserves surface brightness meaning that the relative brightnesses of source and image are characterised entirely by the relative solid angles of source and image, which we have labelled $\vec{\theta_q}$ and $\vec{\theta}$. In general, geometric distortion of an image compared to the source, may be describe by the Jacobian Matrix of the lensing equation (i.e. the rate of change of $\vec{\theta_q}$ with respect to $\vec{\theta}$). Rewriting the Lens Equation (Equation C.3) in vector form using the scalar lensing potential, $\Psi$ (see Equation C.21)

$$\vec{\theta_q} = \vec{\theta} - \frac{D_{ls}}{D_s}\vec{\alpha} = \vec{\theta} - \nabla\Psi \tag{C.23}$$

we can see that the Jacobian matrix, $\frac{\partial\vec{\theta_q}}{\partial\vec{\theta}}$ of the Lens Equation is

$$\frac{\partial\vec{\theta_q}}{\partial\vec{\theta}} = \delta_{ij} - \frac{\partial^2\Psi(\vec{\theta})}{\partial\theta_i\partial\theta_j} \tag{C.24}$$

where $\delta_{ij}$ is the (one-space) Kronecker delta. Now, Poisson's Equation tells us

$$\nabla^2\Psi = \frac{\partial^2\Psi}{\partial^2\theta_1} + \frac{\partial^2\Psi}{\partial^2\theta_2} = 2\kappa \tag{C.25}$$

so, we know that the 11 and 22 terms of the Jacobian Matrix must sum to $2\kappa$. The Jacobian Matrix of the lens mapping can be written

$$\frac{\partial\vec{\theta_q}}{\partial\vec{\theta}} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{C.26}$$

The $\gamma$ terms, which arise from the cross-terms in the lensing potential and are thus due to tidal forces (i.e. as they act with different strength in each of the two basis directions), correspond to how an extended source is distorted into a different shape by lensing (they are *shear* terms). As we always deal with point sources (QSOs) in this thesis, we set $\gamma_1 = \gamma_2 = 0$. The magnification, $\mu$, of a distortion is given by the inverse determinant of the lens mapping

$$\left|\frac{\partial\vec{\theta_q}}{\partial\vec{\theta}}\right|^{-1} = \frac{1}{(1-\kappa)^2} \simeq 1 + 2\kappa \tag{C.27}$$

Where, in the last step we have performed a Taylor Expansion, assuming $\kappa << 1$ (i.e. assuming weak lensing).

### C.3.2 Poisson's Equation and Cosmological Weak Lensing

This section is taken from Bartelmann & Schneider (2001). Further explanation of the theory may be found at that source.

In Newtonian Theory, Poisson's Equation for gravitational fields is

$$\nabla_r^2\Phi = 4\pi G\rho \tag{C.28}$$

On local scales ($<< c/H_o \sim 3000\,h^{-1}$Mpc) and at small peculiar velocites ($<< c$), cosmological inhomogeneites may be described by Newtonian gravity. Expressing the gravitational potential as the sum of a smooth component, $\bar{\Psi}$ and a peturbation due to the inhomogeneity $\Psi$ and similarly expressing the density of the Universe as the sum of the background density and the density of the inhomogeneity $\rho = (1+\delta)\bar{\rho}$, where $\delta$ is the (linear) density contrast (first introduced as Equation B.6) we can consider Poisson's Equation for the inhomogeneity

$$\nabla_r^2 \Phi = 4\pi G \delta\bar{\rho} \tag{C.29}$$

Poisson's Equation describes proper coordinates (see Appendix A). When considering comoving coordinates $\nabla_\chi = \left(\frac{a}{a_o}\right)\nabla_r$ where $a$ is the scale factor of the expansion of the Universe (see Equation A.1) and $\chi$ is the comoving coordinate. In the matter-dominated era of the Universe, the density at any time may be written in terms of the density at the present day, $\rho_o$, as $\rho = \left(\frac{a_o}{a}\right)^3 \rho_o$ (see the discussion after Equation A.17). The critical density of the Universe $\Omega_m$ is defined by Equation A.16. Hence, the cosmological form of Poisson's Equation describing local inhomogeneities in comoving coordinates is given by

$$\nabla_\chi^2 \Phi = 4\pi G \delta\bar{\rho}\left(\frac{a}{a_o}\right)^2 = 4\pi G \delta\bar{\rho}_o\left(\frac{a_o}{a}\right) = \frac{3H_o^2}{2}\frac{a_o}{a}\Omega_m\delta = \frac{3H_o^2}{2}(1+z)\Omega_m\delta \tag{C.30}$$

where $z$ is the redshift of the inhomogeneity.

The reason, in the spherically symmetric case, that the lensing deflection angle, $\alpha$, is given by Equation C.1 is embodied in the Scwharzschild metric, which describes a spherically-symmetric world system with a point-mass at the centre of that system

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right)dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2 \tag{C.31}$$

$$= \left(1 - \frac{\alpha}{2}\right)dt^2 - \left(1 - \frac{\alpha}{2}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2 \tag{C.32}$$

A metric can be constructed that describes an expanding universe pockmarked with inhomogeneities rather than the spherically symmetric smooth space of the Schwarzschild metric (see Bartelmann & Schneider 2001) and that metric is consistent with Equation C.30. In a flat cosmology ($\Omega_m + \Omega_\Lambda = 1$), the form of such a metric implies that the deflection angle, $\alpha$ in the Universe is given by (c.f. Equation C.16)

$$\vec{\alpha}(\vec{\theta}, \chi) = \frac{2}{c^2}\int_0^\chi \frac{\chi - \chi'}{\chi}\nabla_x \Phi(\chi'\vec{\theta}, \chi')d\chi' \tag{C.33}$$

where $\theta$ is, as throughout this appendix, the angle on the sky and $x$ represents directions in the lens plane, often represented by the impact parameter $\xi$ in previous sections ($\xi$ being directions interior to the ray, towards the centre of the Schwarzschild system, rather than any direction in the plane perpendicular to the ray). Thus $\vec{x} = \chi'\vec{\theta}$, where $\chi'$ is the comoving distance from the observer to the lens plane.

In the thin lens approximation (see Equation C.22), the convergence is related to the deflection angle as $\nabla.\left(\frac{D_{ls}}{D_s}\vec{\alpha}(\vec{\theta})\right) = 2\kappa(\vec{\theta})$. In the case of cosmological lensing, a single source is replaced by the mean effect of a large number of sources integrated along the comoving line of sight. For now, we will not consider this component, instead making the convergence and deflection angles functions of the comoving distance $\nabla_\theta.\left(\vec{\alpha}(\vec{\theta},\chi)\right) = \chi'\nabla_x.\vec{\alpha}(\vec{\theta},\chi) = 2\kappa(\vec{\theta},\chi)$. Writing Equation C.33 in terms of the convergence yields what is known as the *effective convergence* for cosmological weak lensing, $\kappa_{eff}(\vec{\theta},\chi)$

$$\kappa_{eff}(\vec{\theta},\chi) = \frac{1}{c^2}\int_0^\chi \frac{\chi-\chi'}{\chi}\chi'\nabla_x^2\Phi(\chi'\vec{\theta},\chi')d\chi' \tag{C.34}$$

Substituting for the appropriate form of Poisson's Equation (Equation C.30) yields

$$\kappa_{eff}(\vec{\theta},\chi) = \frac{3H_o^2\Omega_m}{2c^2}\int_0^\chi \frac{\chi-\chi'}{\chi}\chi'(1+z)\delta(\chi'\vec{\theta},\chi')d\chi' \tag{C.35}$$

So far, we have only considered the foreground contribution of lenses to the effective convergence. However, in the cosmological context, there is also a distribution of sources. If we consider the effective convergence integrated over the number of sources, $N$ in comoving coordinates as

$$\kappa_{eff}(\vec{\theta}) = \int_0^{\chi_H}\frac{dN}{d\chi}\kappa_{eff}(\vec{\theta},\chi)d\chi \tag{C.36}$$

where $\chi_H$ is the integral out to the visible horizon of our Universe, the Equation C.35 can be rewritten, with a little rearranging of the limits of integration, in the practical form

$$\kappa_{eff}(\vec{\theta}) = \frac{3{H_o}^2\Omega_m}{2c^2}\int_o^{\chi_H} d\chi(1+z)\chi\delta(\chi'\theta,\chi)\int_\chi^{\chi_H} d\chi'\frac{\chi'-\chi}{\chi'}\frac{dN}{d\chi'} \tag{C.37}$$

which is used in Chapter 5.

# Bibliography

The following abbreviations are used in this bibliography:

A&A: Astronomy and Astrophysics

A&AS: Astronomy and Astrophysics Supplement

AJ: The Astronomical Journal

ApJ: The Astrophysical Journal

ApJS: The Astrophysical Journal Supplement

Ap&SS: Astrophysics and Space Science

ARA&A: Annual Reviews of Astronomy and Astrophysics

ARep: Astronomy Reports

astro-ph: Astro Preprint Server (http://xxx.soton.ac.uk/archive/astro-ph)

BAAS: Bulletin of the American Astronomical Society

BAN: Bulletins of the Astronomical Institute of the Netherlands

GReGr: General Relativity and Gravitation

Helv. Phys. Acta: Helvetica Physica Acta

INGN: Newsletter of the Isaac Newton Group of Telescopes

IAUS: International Astronomical Union Symposiom

LRR: Living Reviews in Relativity

MNRAS: Monthly Notices of the Royal Astronomical Society

NewAR: New Astronomy Reviews

PASJ: Publications of the Astronomical Society of Japan

PhR: Physics Reports

PhRv: Physical Reviews

PNAS: Proceedings of the National Academy of Sciences of the U.S.A.

RvMP: Reviews of Modern Physics

SPAW: Sitzungsberichte Preußischen Akademie Wißenschaften

Abazajian K., and the Sloan Digital Sky Survey Collaboration, 2003, astro-ph/0305492

Abell G. O., 1958, ApJS, 3, 211

Abell G. O., Corwin H. G., Olowin R. P., 1989, ApJS, 70, 1

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45

Bahcall N. A., Soneira R. M., 1983, ApJ, 270, 20

Bailey J., Glazebrook K., 1999, 2dF User Manual, Anglo-Australian Observatory

Balian R., Schaeffer R., 1989, A&A, 220, 1

Barber A. J., Taylor A. N., 2003, MNRAS, 344, 789

Barber A. J., Thomas P. A., Couchman H. M. P., Fluke C. J., 2000, MNRAS, 319, 267

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barrow J. D., Bhavsar S. P., Sonoda D. H., 1984, MNRAS, 210, 19

Bartelmann M., 1995, A&A, 298, 661

Bartelmann M., 1996, A&A, 313, 697

Bartelmann M., Schneider P., 1993, A&A, 271, 42

Bartelmann M., Schneider P., 1994, A&A, 284, 1

Bartelmann M., Schneider P., 2001, PhR, 340, 291

Bartelmann M., Huss A., Colberg J. M., Jenkins A., Pearce F. R., 1998, A&A, 330, 1

Bartsch A., Schneider P., Bartelmann M., 1997, A&A, 319, 375

Baugh C. M., 1996, MNRAS, 280, 267

Baugh C. M., Efstathiou G., 1993, MNRAS, 265, 145

Baugh C. M., Cole S., Frenk C. S., MNRAS, 283, 1361

Baugh C. M., Benson A. J., Cole S., Frenk C. S., Lacey C. G., 1999, MNRAS, 305, 21

Beck-Winchatz B., Anderson S. F., 1999, AJ, 117, 2582

Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Benítez N., Martínez-González E., 1995, ApJ, 448, 89

Benítez N., Martínez-González E., 1997, ApJ, 477, 27

Benítez N., Sanz J. L., Martínez-González E., 2001, MNRAS, 320, 241

Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587

Bernardeau F., van Waerbeke L., Mellier Y., 1997, A&A, 322, 1

Binney J., Tremaine S., 1987, Galactic Dynamics, Princeton University Press, Chapter 2 (Section 1)

Bischoff O. B., Becker R. H., 1997, AJ, 113, 2000

Blair M., Gilmore G., 1982, PASP, 94, 742

Blandford R., Narayan R., 1986, ApJ, 310, 568

Blanton M., Cen R., Ostriker J. P., Strauss M. A., 1999, ApJ, 522, 590

Blanton M. R., and the Sloan Digital Sky Survey Collaboration, 2001, AJ, 121, 2358

Boissé P., Le Brun V., Bergeron J., Deharveng J. M. 1998, A&A, 333, 841

Bolzonella M., Miralles J.-M., Pelló R., 2000, A&A, 363, 476

Bond J. R., Szalay A., 1983, ApJ, 274, 443

Borgeest U., von Linde J., Refsdal S., 1991, A&A, 251, 35

Bouchet P., Lequeux J., Maurice E., Prevot L., Prevot-Burnichon M. L., 1985, A&A, 149, 330

Boyle B. J., Fong R., Shanks T., 1988, MNRAS, 231, 897

Boyle B. J., Shanks T., Peterson B. A., 1988, MNRAS, 235, 935

Boyle B. J., Jones L. R., Shanks T., 1991, MNRAS, 251, 482

Boyle B. J., Fong R., Shanks T., Peterson B. A., 1987, MNRAS, 227, 717

Boyle B. J., Fong R., Shanks T., Peterson B. A., 1990, MNRAS, 243, 1

Boyle B. J., Shanks T., Croom S. M., Smith R. J., Miller L., Loaring N., Heymans C., 2000, MNRAS, 317, 1014

Boyle B. J., Smith R. J., Shanks T., Croom S. M., Miller L., 1999, IAUS, 183, 178

Braccesi A., Formiggini L., Gandolfi E., 1970, A&A, 5, 264

Brotherton M. S., van Breugel W., Smith R. J., Boyle B. J., Shanks T., Croom S. M., Miller L., Becker R. H., 1998, ApJ, 505, 7

Bukhmastova Y. L., 2001, ARep, 45, 581

Burbidge G., 1979, Nature, 282, 451

Burbidge G., Hoyle F., 1996, A&A, 309, 335

Burbidge G., Hoyle F., Schneider P., 1997, A&A, 320, 8

Burbidge G., Hewitt A., Narlikar J. V., Gupta P. A., 1990, ApJS, 74, 675

Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, ApJ, 429, 582

Canizares C. R., 1981, Nature, 291, 620

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Carlberg R. G., Yee H. K. C., Morris S. L., Lin H., Hall P. B., Patton D., Sawicki M., Shepherd C. W., 2000, ApJ, 542, 57

Carlberg R. G., Yee H. K. C., Morris S. L., Lin H., Hall P. B., Patton D. R., Sawicki M., Shepherd C. W., 2001, ApJ, 552, 427

Colin P., Klypin A. A., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 523, 32

Cole S., Hatton S., Weinberg D. H., Frenk C. S., 1998, MNRAS, 300, 945

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168

Colless M., 1998, in Wide Field Surveys in Cosmology, Editions Frontieres, ISBN 2-863-32241-9, 77

Colless M., and the 2dF Galaxy Redshift Survey Collaboration, 2001, MNRAS, 328, 1039

Colless M., and the 2dF Galaxy Redshift Survey Collaboration, 2003, astro-ph/0306581

Colley W. N., Gott III J. R., Weinberg D. H., Park C., Berlind A. A., 2000, ApJ, 529, 795

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693

Connolly A. J., and the Sloan Digital Sky Survey Collaboration, 2002, ApJ, 579, 42

Cooray A., Sheth R., 2002, PhR, 372, 1

Corradi R. L. M., Dee K. M., Bassom R. A., Blanken M. F., Goodsell S. J., van der Hoeven M., 2001, INGN, 5, 19

Crampton D., Cowley A. P., Hartwick F. D. A., 1987, ApJ, 314, 129

Croom S. M., 1997, PhD Thesis, University of Durham

Croom S. M., Shanks T., 1996, MNRAS, 281, 893

Croom S. M., Shanks T., 1999, MNRAS, 307, 17

Croom S. M., Smith R. J., Boyle B. J., Shanks T., Loaring N. S., Miller L., Lewis I. J., 2001, MNRAS, 322, 29

Croom S. M., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2003, MNRAS, submitted

Dalton G. B., Maddox S. J., Sutherland W. J., Efstathiou G., 1997, MNRAS, 289, 263

Davis M., Efstathiou G., Frenk C., White S. D. M., 1985, ApJ, 292, 371

de Sitter W., 1916, MNRAS, 76, 699

Debes J. H., Sigurdsson S., 2002, ApJ, 572, 556

Demianski M., de Ritis R., Marino A. A., Piedipalumbo E., 2000, astro-ph/0004376

Diaferio A., Kauffmann G., Colberg J. M., White S. D. M., 1999, MNRAS, 307, 537

Dodelson S., and the Sloan Digital Sky Survey Collaboration, 2002, ApJ, 572, 140

Dolag K., Bartelmann M., 1997, MNRAS, 291, 446

Drinkwater M. J., Phillipps S., Jones J. B., Gregg M. D., Deady J. H., Davies J. I., Parker Q. A., Sadler E. M., Smith R. M., 2000, A&A, 355, 900

Dyer C. C., Roeder R. C., 1974, ApJ, 189, 167

Dyer C. C., Roeder R. C., 1981, GReGr, 13, 1157

Efstathiou G., Bernstein G., Tyson J. A., Katz N., Guhathakurta P., 1991, ApJ, 380, 47

Efstathiou G., and the 2dF Galaxy Redshift Survey Collaboration, 2002, MNRAS, 330, 29

Einstein A., Über den Einflußder Schwerkraft auf die Ausbreitung des Lichtes, 1911, Annalen der Physik, 35, 898

Einstein A., Erklärungder Perihelbewegung des Merkur aus der allgemeinen Relitivitätstheorie, 1915, SPAW, 831

Einstein A., Die Grundläge der allgemeinen Relitivitätstheorie, 1916, Annalen der Physik, 49, 769

Einstein A., 1917, SPAW, 142

Eke V., et al., in preparation.

Ellison S. L., Pettini M., Steidel C. C., Shapley A. E., 2001, AJ, 549, 770

Evrard A. E., and the Virgo Consortium, 2002, ApJ, 573, 7

Fall M. S., 1979, RvMP, 51, 21

Fan X., and the Sloan Digital Sky Survey Collaboration, 2001, AJ, 121, 54

Ferguson H. C., 1993, MNRAS, 263, 343

Ferreras I., Benítez N., Martínez-González E., 1997, AJ, 114, 1728

Fitzpatrick E. L., Massa D., 1990, ApJS, 72, 163

Freedman W. L., and the Hubble Space Telescope Key Project Collaboration, 2001, ApJ, 553, 47

Friedmann A., 1922, Zeitschrift fur Physik, 10

Fry J. N., 1986, ApJ, 308, 71

Fry J. N., Gaztañaga E., 1993, ApJ, 413, 447

Fugmann W., 1988, A&A, 204, 73

Fugmann W., 1990, A&A, 240, 11

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ 111, 1748

Garcia A. M., 1993, A&AS, 100, 47

Gaztañaga E., 1994, MNRAS, 268, 913

Gaztañaga E., 1995, ApJ, 454, 561

Gaztañaga E., 2003, ApJ, 589, 82

Gaztañaga E., Yokoyama J., 1993, ApJ, 403, 450

Geller M. J., Huchra J. P., 1983, ApJS, 52, 61

Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 2000, ApJ, 544, 616

Gilmore D. M., Allen R. J., 1990, BAAS, 22, 1327

Girardi M., Mezzetti M., Giurcin G., Mardirossian F., 1992, ApJ, 394, 442

Gourgoulhon E., Chamaraux P., Fouque P., 1992, A&A, 255, 69

Graham M. J., Clowes R. G., Campusano L. E., 1999, ApJ, 513, 69

Griffiths L. M., Melchiorri A., Silk J., 2001, ApJ, 553, 5

Groth M. G., Peebles P. J. E., 1977, ApJ, 217, 385

Guimãraes A. C. C., 2002, MNRAS 337, 631

Guth A. H., 1981, PhRvD, 23, 347

Hall P. B., Osmer P. S., Green R. F., Porter A. C., Warren S. J., 1996, ApJ, 462, 614

Hall P. B., and the CNOC2 collaboration, 2000, AJ, 120, 2220

Hamilton A. J. S., 1993, ApJ, 417, 19

Hamilton A. J. S., Kumar P., Lu E., Matthews A., 1991, ApJ, 374, 1

Hartwick F. D. A., Schade D., 1990, ARA&A, 28, 437

Hauser M. G., Peebles P. J. E., 1973, ApJ, 185, 757

Hewitt J. N., Turner E. L., Schneider D. P., Burke B. F., Langston G. I., Lawrence C. R., 1988, Nature, 333, 537

Hills J. G., 1981, AJ, 86, 1730

Hoekstra H., Franx M., Kuijken K., Carlberg R. G., Yee H. K. C., Lin H., Morris S. L., Hall P. B., Patton D. R., Sawicki M., Wirth G. D., 2001, ApJ, 548, 5

Hoekstra H., van Waerbeke L., Gladders M. D., Mellier Y., Yee H. K. C., 2002, ApJ, 577, 604

Hubble E., 1929, PNAS, 15, 168

Hubble E., 1934, ApJ, 79, 8

Huchra J. P., Geller M. J., ApJ, 257, 423

Jenkins A., Pearce F. R., Thomas P. A., Colberg J. M., White S. D. M., Couchman H. M. P., Peacock J. A., Efstathiou G., Nelson A. H., 1998, ApJ, 499, 20

Kaiser N., 1984, ApJ, 284, 9

Kantowski R., Kao J. K., Thomas R. C., 2000, ApJ, 545, 549

Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188

Kayser R., Helbig P., Schramm T., 1997, A&A, 318, 680

Kennefick J. D., Osmer P. S., Hall P. B., Green R. F., 1997, AJ, 114, 2269

Kirshner R. P., 1977, ApJ, 212, 319

Kinney A. L., Calzetti D., Bica E., Storchi-Bergmann T., 1994, ApJ, 429, 172

Kippenhahn R., DeVries H., 1974, Ap&SS, 26, 131

Kneib J. P., 1993, PhD Thesis, Université Paul Sabatier, Toulouse

Kochanek C. S., Falco E. E., Impey C. D., Lehár J., McLeod B. A., Rix H. W., the CASTLES collaboration, 2002, `http://cfa-www.harvard.edu/castles/`

Komatsu E., and the WMAP Collaboration, 2003, astro-ph/0302223

Koo D. C., Kron R. G., 1988, ApJ, 325, 92

Koo D. C., Kron R. G., Cudworth K. M., 1986, PASP, 98, 285

Koornneef J., 1982, A&A, 107, 247

Kovner I., 1989, ApJ, 341, 1

Lahav O., and the 2dF Galaxy Redshift Survey Collaboration, 2002, MNRAS, 335, 432

Landy S. D., Szalay A. S., ApJ, 412, 64

Lebach D. E., Corey B. E., Shapiro I. I., Ratner M. I., Webber J. C., Rogers A. E. E., Davis J. L., Herring T. A., 1995, PhRvL, 75, 1439

Lemaitre G., 1927, Annales Societe Scientifique Bruxelles, 47

Lewis I. J., et al., 2002, MNRAS, 333, 279

Liebert J., Dahn C. C., Monet D. G., 1988, ApJ, 332, 891

Limber D. N., 1953, ApJ, 117, 134

Lin H., Kirschner P., Shectman S. A., Landy S. D., Oemler A., Tucker D. L., Schechter P. L., 1996, ApJ, 471, 617

Loveday J., Efstathiou G., Peterson B. A., Maddox S. J., 1992, ApJ, 400, 43

Loveday J., Peterson B. A., Efstathiou G., Maddox S. J., 1992, ApJ, 390, 338

Lynds R., Petrosian V., 1986, BAAS, 18, 1014

Maddox S. J., Efstathiou G., Sutherland W. J., 1990, MNRAS, 246, 433

Maddox S. J., Efstathiou G., Sutherland W. J., 1996, MNRAS, 283, 1227

Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1990, MNRAS, 243, 692

Madgwick D. S., Hewett P. C., Mortlock D. J., Lahav O., 2002, MNRAS, 334, 209

Marano B., Zamorani G., Zitelli V., 1988, MNRAS, 232, 111

Marshall H. L., Huchra J. P., Tananbaum H., Avni Y., Braccesi A., Zitelli V., Zamorani G., 1984, ApJ, 283, 50

Metcalfe N., Fong R., Shanks T., 1995, MNRAS, 274, 769

Misner C. W., Thorne K. S., Wheeler J. A., 1973, in *Gravitation*, W H Freeman & Co., ISBN 0-716-70344-0

Monier E. M., Kennefick J. D., Hall P. B., Osmer P. S., Smith M. G., Dalton G. B., Green R. F., 2002, AJ, 124, 2971

Moore B., Quinn T., Governato F., Stadel J., Lake G., 1999, MNRAS, 310, 1147

Maoz D., 1995, ApJ, 455, 115

Maoz D., Rix H. W., Gal-Yam A., Gould A., 1997, AJ, 486, 75

Martinez H. J., Merchán M. E., Valotto C. A., Lambas D. G., 1999, ApJ, 514, 558

Menárd B., Bartelmann M., 2002, A&A, 386, 784

Merchán M. E., Zandivarez A., 2002, MNRAS, 335, 216

Merchán M. E., Maia M. A. G., Lambas D. G., 2000, ApJ, 545, 26

Metcalfe N., Fong R., Shanks T., 1995, MNRAS, 274, 769

Miller L., Lopes A. M., Smith R. J., Croom S., Boyle B. J., Shanks T., Outram P., astro-ph/0210644

Miralles J. M., Erben T., Hämmerle H., Schneider P., Fosbury R. A. E., Freudling W., Pirzkal N., Jain B., White S. D. M., 2002, A&A, 388, 68

Muñoz-Tuñòn C., Varela A. M., Mahoney T., 1998, NewAR, 42, 409

Narayan R., 1989, ApJ, 339, 53

Navarro J. F., Frenk C. S., White S. D. M., 1995, ApJ, 275, 720

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, AJ, 486, 75

Nollenberg J. G., Williams L. L. R., Maddox S. J., 2003, AJ, 125, 2927

Norberg P., 2003, PhD Thesis, University of Durham

Norberg P., and the 2dF Galaxy Redshift Survey Collaboration, 2002, MNRAS, 332, 827

Norberg P., and the 2dF Galaxy Redshift Survey Collaboration, 2002, MNRAS, 2002, MNRAS, 336, 907

Norman D. J., Impey C. D., 1999, AJ, 118, 613

Norman D. J., Williams L. L. R., 2000, AJ, 118, 613

Norman D. J., Impey C. D., 2001, AJ, 121, 2392

O'Donnell J. E., 1994, ApJ, 422, 158

Okamoto T., Hu W., 2003, PhRvD, 67, 3002

Oort J. H., 1950, BAN, 11, 91

Osmer P. S., Smith M. G., 1980, ApJS, 42, 333

Outram P. J., Hoyle F., Shanks T., Croom S. M., Boyle B. J., Miller L., Smith R. J., Myers A. D., 2003, MNRAS, 342, 483

Paling S., 2003, IAUS, 220, 203

Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020

Peacock J. A., Dodds S. J., 1996, MNRAS, 280, 19

Peacock J. A., and the 2dF Galaxy Redshift Survey Collaboration, 2001, Nature, 410, 169

Peebles P. J. E., 1966, ApJ, 146, 542

Peebles P. J. E., 1973, ApJ, 185, 413

Peebles P. J. E., 1980, in The Large Scale Structure in the Universe, Princeton University Press, ISBN 0-691-08240-5

Peebles P. J. E., 1982, ApJ, 258, 415

Peebles P. J. E., 1993, in Principles of Physical Cosmology, Princeton University Press, ISBN 0-691-01933-9

Peebles P. J. E., Hauser M. G., 1974, ApJS, 28, 19

Peebles P. J. E., Groth M. G., 1976, A&A, 53, 131

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 421

Percival W. J., and the 2dF Galaxy Redshift Survey Collaboration, 2001, MNRAS, 327, 1297

Perlmutter S., and the Supernova Cosmology Project Collaboration, 1999, ApJ, 517, 565

Phillipps S., Fong R., Fall R. S., Ellis S. M., MacGillivray H. T., 1978 MNRAS, 182, 673

Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezic Z., 2003, AJ, 125, 1559

Plionis M., Basilakos S., 2002, MNRAS, 329, 47

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14

Ramella M., Geller M. J., Pisani A., da Costa L. N., 2002, AJ, 123, 2976

Reblinsky K., Bartelmann M., 1999, A&A, 345, 1

Rengstorf A. W., and the QUEST Collaboration, 2001, BAAS, 199, 13807

Richards G. T., and the Sloan Digital Sky Survey Collaboration, 2002, AJ, 123, 2945

Robertson H. P., 1935, ApJ, 82, 284

Rodrigues-Williams L. L., Hogan C. J., 1994, AJ, 107, 451

Rubin V. C., D'Odorico S., 1969, A&A, 2, 484

Rubin V. C., Burstein D., Ford Jr. W. K., Thonnard N., 1985, ApJ, 289, 81

Sand D. J., Treu T., Ellis R., 2002, ApJ, 574, 129

Santiago B. X., Strauss M. A., Lahav O., Davis M., Dressler A., Huchra J. P. 1995, ApJ, 446, 457

Sanz J. L., Martínez-González E., Benítez N., 1997, MNRAS, 291, 418

Schade D., Crampton D., Hammer F., Le Fevre O., Lilly S. J., 1996, MNRAS, 278, 95

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Schmidt M., 1963, Nature, 197, 1040

Schmidt M., 1965, ApJ, 141, 1

Schmidt M., 1968, ApJ, 151, 393

Schmidt M., 1976, ApJ, 209, 55

Schmidt M., Green R. F., 1983, ApJ, 269, 352

Schneider D. P., and the Sloan Digital Sky Survey Collaboration, 2002, AJ, 123, 567

Schneider P., 1989, A&A, 221, 221

Schneider P., Ehlers J., Falco E. E., 1992, in Gravitational Lenses, Springer-Verlag, ISBN 0-387-97070-3

Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629

Scranton R., and the Sloan Digital Sky Survey Collaboration, 2002 ApJ, 579, 48

Seldner M., Peebles P. J. E., 1977, ApJ, 215, 703

Seldner M., Peebles P. J. E., 1979, ApJ, 227, 30

Seljak U., 2000, MNRAS, 318, 203

Shanks T., Boyle B. J., 1994, MNRAS, 271, 753

Shanks T., Fong R., Green M. R., Clowes R. G., Savage A., 1983, MNRAS, 203, 181

Sharp R. G., McMahon R. G., Irwin M. J., Hodgkin S. T., 2001, MNRAS, 326, 45

Sheth R. K., Diaferio A., 2001, MNRAS, 322, 901

Sion E. M., 1984, ApJ, 282, 612

Smith J., and the Sloan Digital Sky Survey Collaboration, 2002, AJ, 123, 2121

Smith R. E., Peacock J. A., Jenkins A., White S. D. M., Frenk C. S., Pearce F. R., Thomas P. A., Efstathiou G., Couchman H. M. P., 2003, MNRAS, 341, 1311

Smith R. J., Boyle B. J., Shanks T., Croom S. M., Miller L., Read M., 1997, IAUS, 179, 348

Smoot G. F., and the COBE collaboration, 1992, ApJ, 396, 1

Soucail G., Fort B., Mellier Y., Picat J. P., 1987, A&A, 172, 14

Spergel D. N., and the WMAP Collaboration, 2003, astro-ph/0302209

Stevenson P. R. F., Fong R., Shanks T., 1988, MNRAS, 234, 801

Stevenson P. R. F., Shanks T., Fong R., MacGillivray H. T., 1985, MNRAS, 213, 953

Storrie-Lombardi L. J., Irwin M. J., McMahon R. G., Hook I. M., 2001, MNRAS, 322, 933

Stoughton C., and the Sloan Digital Sky Survey Collaboration, 2002, AJ, 123, 485

Strauss M. A., Ostriker J. P., Cen R., 1998, ApJ, 494, 20

Sugiyama N., 1995, ApJS, 100, 281

Tadros H., Efstathiou G., 1996, MNRAS, 282, 1381

Thomas P. A., Webster R. L., Drinkwater M. J., 1995, MNRAS, 273, 1069

Thomas P. A., Colberg J. M., Couchman H. M. P., Efstathiou G. P., Frenk C. S., Jenkins A. R., Nelson A. H., Hutchings R. M., Peacock J. A., Pearce F. R., White S. D. M., 1998, MNRAS, 296, 1061

Totsuji H., Kihara T., 1969, PASJ, 21, 221

Tully R. B., 1987, ApJ, 321, 280

Turner E. L., Gott J. R., 1976, ApJ, 32, 409

Turner E. L., Ostriker J. P., Gott III J. R., 1984, ApJ, 284, 1

Tyson J. A., 1986, AJ, 92, 691

Tyson J. A., Valdes F., Wenk R. A., 1990, ApJL, 349, 1

Verde L., and the 2dF Galaxy Redshift Survey Collaboration, 2002, MNRAS, 335, 432

Véron P., 1983, in Quasars and Gravitational Lenses, Proceedings of the 24th Liege International Astrophysical Colloquium, 210

Véron-Cetty M. P., Véron P., 2001, A&A, 374, 92

Vielva P., Mart'ınez-González E., Gallegos J. E., Toffolatti L., Sanz J. L., 2003, MNRAS, 344, 89

Vikhlinin A., McNamara B. R., Hornstrup A., Quintana H., Forman W., Jones C., Way M., 1999, ApJ, 520, 1

Wadadekar Y., Kembhavi A., 1999, AJ, 118, 1435

Walker A. G., 1936, Procedures of the London Mathematical Society, 2nd Series, 42

Walsh D., Carswell R. F., Weymann R. J., 1979, Nature, 279, 381

Wambsganss J., 1998, LRR 1, 12, http://www.livingreviews.org/Articles/Volume1/1998-12wamb/

Warren S. J., Hewett P. C., Osmer P. S., 1994, ApJ, 421, 412

Webster R. L., Hewett P. C., Harding M. E., Wegner G. A., 1988, Nature, 336, 558

White M., Kochanek C. S., 2001, ApJ, 574, 24

White M., Van Waerbeke L., Mackey J., 2002, ApJ, 575, 640

White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nature, 366, 429

Williams L. L. R., Irwin M., 1998, MNRAS, 298, 378

Williams L. L. R., Frey N., 2003, ApJ, 583, 594

Willmer C. N. A., da Costa L. N., Pellegrini P. S., 1998, AJ, 115, 869

Wilson G., 2003, ApJ, 585, 191

Wu X. P., 1994, A&A, 286, 748

Wu X. P., Fang L. Z., Zhu Z., Qin B., 1996, ApJ, 471, 575

Wyithe J. S. B., Winn J. N., Rusin D., 2003, ApJ, 583, 58

Yasuda N., and the Sloan Digital Sky Survey Collaboration, 2001, AJ, 122, 1104

York D. G., and the Sloan Digital Sky Survey Collaboration, 2000, AJ, 120, 1579

Zehavi I., and the Sloan Digital Sky Survey Collaboration, 2002, ApJ, 571, 172

Zitelli V., Mignoli M., Zamorani G., Marano B., Boyle B. J., 1992, MNRAS, 256, 349

Zwicky F., 1933, Helv. Phys. Acta, 6, 110